

Zusammenhänge zwischen der Beschulungsart und den schulischen Kompetenzen sowie der schulischen Motivation von Kindern mit sonderpädagogischem Förderbedarf

Dissertation

zur Erlangung des akademischen Grads

Doctor rerum naturalium (Dr. rer. nat.)

im Fach Psychologie

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von Dipl.-Psych. Aleksander Kocaj

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachterinnen & Gutachter

1. Prof. Dr. Petra Stanat
2. Prof. Dr. Manuel Völkle
3. Prof. Dr. Kai Schnabel Cortina

Tag der Verteidigung: 05. März 2018

Inhaltsverzeichnis

Zusammenfassung	7
Abstract	9
Liste der Beiträge	11
1. Einleitung	13
2. Theoretischer Hintergrund	19
2.1 Historischer Überblick über das sonderpädagogische Fördersystem in Deutschland.....	19
2.2 Inklusionsdefinitionen	21
2.2.1 Definition von Inklusion in der UN-Behindertenrechtskonvention	21
2.2.2 Definitionen von Inklusion in der empirischen Bildungsforschung.....	23
2.2.3 Abgrenzung der Begriffe Inklusion und Integration	25
2.2.4 Inklusionsdefinition der vorliegenden Dissertation.....	29
2.3 Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf.....	30
2.3.1 Definition sonderpädagogischer Förderbedarf	30
2.3.2 Verfahren der Feststellung eines sonderpädagogischen Förderbedarfs.....	36
2.3.3 Situation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Deutschland.....	42
2.4 Methodische Herausforderungen beim Einbezug von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments	52
2.4.1 Kompetenztestung bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments	53
2.4.2 Herausforderungen bei der Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf	55
2.4.3 Umgang mit methodischen Herausforderungen bei der Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf	58
2.4.4 Forschungsstand zur Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments	60
2.5 Der Zusammenhang zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf.....	63
2.5.1 Rahmenmodell zum Zusammenhang von schulischer und familiärer Lernumgebung mit schulischen Outcomes	63
2.5.2 Förderschulen und allgemeine Schulen als differenzielle Lern- und Entwicklungsmilieus.....	65
2.5.3 Schulische Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen.....	79
2.5.4 Schulische Motivation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen.....	86
3. Forschungsfragen und Datengrundlage der Dissertation.....	99
3.1 Forschungsstand und Forschungsbedarf zum Zusammenhang zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit SPF.....	99

3.2 Datengrundlage der Dissertation.....	102
3.3 Forschungsfragen und Teilstudien.....	104
4. Teilstudien der Dissertation.....	109
4.1 Teilstudie 1: Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf.....	109
4.2 Teilstudie 2: Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe	134
4.3 Teilstudie 3: Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf.....	166
4.4 Teilstudie 4: Educational Placement and Achievement Motivation of Students With Special Educational Needs	197
5. Diskussion.....	228
5.1 Zusammenfassung und Interpretation der Befunde	228
5.1.1 Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (Teilstudie 1).....	228
5.1.2 Vergleich der schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen (Teilstudie 2).....	231
5.1.3 Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf (Teilstudie 3).....	233
5.1.4 Vergleich der schulischen Motivation von Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen (Teilstudie 4).....	235
5.2 Inhaltliche und methodische Herausforderungen bei der Interpretation der vorliegenden Ergebnisse	238
5.2.1 Auswirkungen der Datengrundlage auf die Interpretation der vorliegenden Ergebnisse.....	238
5.2.2 Methodische Herausforderungen bei der Schätzung von Schularteffekten.....	240
5.2.3 Statistische Kontrolle von Selektionseffekten durch Propensity-Score-Matching.....	242
5.3 Diskussion der Zusammenhänge zwischen Beschulungsart und schulischen Outcomes von Kindern mit sonderpädagogischem Förderbedarf.....	243
5.4 Implikationen für zukünftige Forschung	248
5.4.1 Forschungsbedarfe für zukünftige Studien.....	249
5.4.2 Forschungsausblick auf zwei Anschlussprojekte	250
5.5 Implikationen für die Bildungspolitik und pädagogische Praxis	254
5.5.1 Implikationen für die Bildungspolitik	254
5.5.2 Implikationen für die pädagogische Praxis	257
5.6 Fazit	260
Literaturverzeichnis.....	262
Danksagung.....	284
Anhang.....	285

Abbildungsverzeichnis

<i>Abbildung 1.</i> Prozentualer Anteil der Schülerinnen und Schüler mit den Förderschwerpunkten Lernen, Sprache und emotionale und soziale Entwicklung an allen Schülerinnen und Schülern mit SPF	44
<i>Abbildung 2.</i> Entwicklung der Inklusionsquote nach Förderschwerpunkt vom Schuljahr 1999/2000 bis zum Schuljahr 2014/2015	45
<i>Abbildung 3.</i> Sonderpädagogische Förderung im Schuljahr 2014/2015 in den Bundesländern	47
<i>Abbildung 4.</i> Kontext-Input-Prozess-Output-Modell zur Untersuchung von Effekten der Schule und des Unterrichts auf Schülerinnen und Schüler.....	64
<i>Abbildung 5.</i> Vermittlungsmodell von Effekten der Komposition und Institution auf die individuelle schulische Entwicklung.....	73
<i>Abbildung 6.</i> Der Big-Fish-Little-Pond-Effekt.	90
<i>Abbildung 7.</i> Standardisierte Mittelwertunterschiede zwischen Kindern mit SPF in Regel- und Förderschulen vor und nach dem Full Matching.	152
<i>Abbildung 8.</i> Zusammenhang zwischen sozialer Zusammensetzung (HISEI) und mittlerem Niveau der nonverbalen kognitiven Grundfähigkeiten (KFT) auf Klassenebene (linearer und quadratischer Trend). .	184

Tabellenverzeichnis

<i>Tabelle 1.</i> Vergleich zwischen inklusiven und integrativen Praktiken nach Hinz	27
<i>Tabelle 2.</i> Stichprobenumfänge in den drei untersuchten Schülergruppen	119
<i>Tabelle 3.</i> Relativer Anteil fehlender Werte pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen.....	122
<i>Tabelle 4.</i> Itemfit pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen	124
<i>Tabelle 5.</i> Differenzielle Itemfunktionen pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen.....	125
<i>Tabelle 6.</i> Korrelationen zwischen schulischen Kompetenzen und kognitiven Grundfähigkeiten in den drei untersuchten Schülergruppen.....	128
<i>Tabelle 7.</i> Stichprobenumfänge im Überblick.....	141
<i>Tabelle 8.</i> Unterschiede zwischen Kindern mit SPF nach Art der Beschulung	147
<i>Tabelle 9.</i> Ergebnisse der logistischen Regression zur Vorhersage des Besuchs einer Regelschule	149
<i>Tabelle 10.</i> Verteilung der Propensity Scores vor und nach dem Matching und standardisierter Bias.....	151
<i>Tabelle 11.</i> Effekte des Besuchs einer Regelschule auf die Kompetenzen in Deutsch und Mathematik nach dem Full Matching.....	153
<i>Tabelle 12.</i> Stichprobengröße nach Beschulungsart.....	176
<i>Tabelle 13.</i> Klassengröße nach Beschulungsart	176
<i>Tabelle 14.</i> Demografische und sozio-kulturelle Unterschiede zwischen Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen: metrisch skalierte Merkmale	181
<i>Tabelle 15.</i> Demografische und sozio-kulturelle Unterschiede zwischen Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen: nominal skalierte Merkmale	182
<i>Tabelle 16.</i> Korrelationen zwischen Art der Beschulung und Merkmalen der Klassenzusammensetzung.	185
<i>Tabelle 17.</i> Deskriptive Kennwerte der mittleren Klassenzusammensetzung in Förder- und Grundschulen	185
<i>Tabelle 18.</i> Mehrebenenmodell zur Vorhersage der Mathematikleistung durch individuelle Schülermerkmale, Beschulungsart und Klassen-komposition.....	188
<i>Tabelle 19.</i> Mehrebenenmodell zur Vorhersage der Leseleistung durch individuelle Schülermerkmale, Beschulungsart und Klassenkomposition	189
<i>Tabelle 20.</i> Sample Sizes and Descriptive Statistics by Educational Placement.....	206
<i>Tabelle 21.</i> Correlations Between Educational Placement, Achievement, and Social Support	211
<i>Tabelle 22.</i> Predicting SEN Students' Academic Self-Concept in German.....	213
<i>Tabelle 23.</i> Predicting SEN Students' Academic Self-Concept in Mathematics.....	215
<i>Tabelle 24.</i> Predicting SEN Students' Enjoyment of Learning With Reading Achievement	216

Zusammenfassung

Im Zuge der Umsetzung der UN-Behindertenrechtskonvention (2008) werden Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) verstärkt in allgemeinen Schulen¹ anstatt in separaten Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016; Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, KMK, 2016). Diese Umstellung im deutschen Bildungssystem hin zum gemeinsamen Unterricht von Schülerinnen und Schülern mit und ohne SPF ist eine der umfangreichsten Reformen in den letzten 20 Jahren (Baumert, 2016a; Weishaupt, 2016). Es besteht aber Forschungsbedarf, wie die Beschulungsart mit Merkmalen der schulischen Entwicklung von Kindern und Jugendlichen mit SPF zusammenhängt. Dieser Frage wird in der vorliegenden Dissertation in vier Teilstudien nachgegangen. Grundlage der Dissertation sind Daten des IQB-Ländervergleichs 2011 (Stanat, Pant, Böhme & Richter, 2012), ein bundesweites Large-Scale Assessment am Ende der 4. Jahrgangsstufe.

In der ersten Teilstudie prüften wir, ob die Kompetenztests im IQB-Ländervergleich 2011 geeignet sind, die schulischen Kompetenzen von Kindern mit SPF in Förderschulen und allgemeinen Schulen reliabel und valide zu erfassen. Dazu wurden der Anteil fehlender Werte, die Passung des Raschmodells, differenzielle Itemfunktionen und die Korrelationen der Testwerte zwischen den Schülergruppen mit SPF in Förderschulen und allgemeinen Schulen sowie Schülerinnen und Schülern ohne SPF analysiert. Insgesamt waren die Kompetenztests geeignet, die schulischen Kompetenzen von Kindern mit SPF reliabel und valide zu erfassen. Die Messeigenschaften der Kompetenztests waren zwischen den Schülergruppen mit SPF in Förderschulen und allgemeinen Schulen vergleichbar. Jedoch zeigten sich Einschränkungen in der Testgüte für die Kompetenztests im Fach Mathematik für Kinder mit SPF in Förderschulen.

In der zweiten Teilstudie verglichen wir die schulischen Kompetenzen in den Fächern Deutsch (Lesen und Zuhören) und Mathematik zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen. Bei der Auswertung kamen *Propensity-Score-Matching*-Verfahren zum Einsatz, um potenzielle Unterschiede in Lernvoraussetzungen und weiteren Hinter-

¹ In dieser Dissertation werden unter allgemeinen Schulen allgemeinbildende Schularten ohne Förderschulen verstanden (KMK, 2015). Die Begriffe *Regelschulen* und *Grundschulen* werden dabei synonym mit dem Begriff *allgemeine Schulen* gebraucht.

grundmerkmalen zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen statistisch zu kontrollieren. Kinder mit SPF in allgemeinen Schulen erzielten in beiden Fächern höhere schulische Kompetenzen als Kinder mit SPF in Förderschulen mit vergleichbaren Lernvoraussetzungen und Hintergrundmerkmalen. Insbesondere Kinder mit dem Förderschwerpunkt *Lernen* profitierten von einer Beschulung in allgemeinen Schulen. Kinder mit dem Förderschwerpunkt *Sprache* in allgemeinen Schulen wiesen ebenfalls höhere schulische Kompetenzen im Lesen und Zuhören auf als vergleichbare Kinder mit SPF in Förderschulen, die Schulartunterschiede waren aber geringer ausgeprägt und im Fach Mathematik nicht statistisch signifikant.

Die dritte Teilstudie schließt an den Ergebnissen der zweiten Teilstudie an und untersuchte, ob die Klassenkomposition in Förderschulen und allgemeinen Schulen mit den schulischen Kompetenzen von Kindern mit SPF zusammenhängt. Die fähigkeitsbezogene Klassenkomposition, jedoch nicht die soziale Klassenzusammensetzung, hing unter Berücksichtigung individueller Hintergrundmerkmale und Schulart positiv mit den schulischen Kompetenzen von Kindern mit SPF zusammen. Kinder mit SPF wiesen in Schulklassen mit einem höheren Fähigkeitsniveau (erfasst über die kognitiven Grundfähigkeiten) auch höhere schulische Kompetenzen in den Fächern Deutsch und Mathematik auf. Diese Ergebnisse weisen darauf hin, dass die Klassenkomposition zu Schulartunterschieden in den schulischen Kompetenzen von Kindern mit SPF beiträgt.

In der vierten Teilstudie untersuchten wir Zusammenhänge zwischen der Beschulungsart und Merkmalen der schulischen Motivation. Kinder mit SPF in Förderschulen wiesen ein höheres akademisches Selbstkonzept in den Fächern Deutsch und Mathematik sowie fächerübergreifend eine höhere Lernfreude auf als Kinder mit SPF in allgemeinen Schulen. Diese motivationalen Unterschiede konnten auf die geringere Leistungsstärke in Förderschulen im Vergleich zu allgemeinen Schulen zurückgeführt werden. Die Ergebnisse stehen in Einklang mit dem sogenannten Fischteicheffekt (Marsh, 1987), wonach Schülerinnen und Schüler soziale Vergleiche bei der Ausbildung ihres akademischen Selbstkonzepts und ihrer Lernfreude nutzen. Außerdem berichteten Kinder mit SPF, die sich von ihren Mitschülerinnen und Mitschülern emotional unterstützt fühlten, eine höhere schulische Motivation.

Die zentralen Ergebnisse der Dissertation weisen somit auf gegenläufige Zusammenhänge zwischen der Beschulungsart und den schulischen Kompetenzen beziehungsweise der schulischen Motivation von Kindern mit SPF hin, die bei der Unterrichtsgestaltung berücksichtigt werden müssen.

Abstract

With the ratification of the UN Convention on the rights of persons with disabilities (2008), Germany has committed to implement a more inclusive school system and to increase the proportion of students with special educational needs (SEN) in regular schools. This change in the German education system is one of the largest educational reforms in the last 20 years because the majority of SEN students in Germany still receives instruction in special education schools (Baumert, 2016a). The purpose of this dissertation is to investigate the relationship between the educational placement of students with SEN in special education schools or regular schools and cognitive and motivational aspects of their school development. The dissertation consists of four studies that are based on data from the *National Assessment Study 2011* (IQB-Ländervergleich 2011; Stanat, Pant, Böhme & Richter, 2012), a nationally representative cross-sectional large-scale assessment of German fourth-graders.

In the first study, we investigated if the achievement tests in German and mathematics assess SEN students' proficiencies reliably and validly. Furthermore, we examined if the interpretation of the test scores are comparable between students with SEN in special education schools or regular schools and students without SEN. To evaluate these aspects of the achievement tests' structural validity, we analyzed missing patterns, fit of the Rasch model, differential item functioning, and correlations of the test scores. Overall, the achievement tests captured students' proficiencies in similar ways for SEN students in special education schools or regular schools and students without SEN. The interpretation of the test scores was also comparable between students with SEN in special education schools and regular schools. The comparability of the achievement tests in mathematics, however, was constrained for SEN students in special education schools compared to students without SEN.

In the second study, we compared SEN students' school achievement in German (reading listening comprehension) and mathematics between special education schools and regular schools. As students cannot be randomly assigned to different school types, we applied propensity score matching to control for confounding variables (i.e., differences in achievement-related characteristics). After matching, SEN students in regular schools had significantly higher test scores across all domains than comparable SEN students in special education schools. In particular, students with special needs in learning benefitted from an education in regular schools. SEN students with special needs in speech and language in regular schools also showed higher test scores than comparable SEN students in special education schools.

The differences between school types, however, were smaller in German and statistically not significant in mathematics.

The third study examined the relation between classroom composition and SEN students' school achievement in special education schools and regular schools. We found a significant positive relationship between the average cognitive abilities at the class level and SEN students' individual test scores in German and mathematics. In contrast, the socioeconomic classroom composition was not significantly related to SEN students' individual achievement after accounting for students' achievement-related characteristics and school type. These results indicate that compositional effects could contribute to achievement differences between SEN students in special education schools and regular schools.

The fourth study focused on the relation between SEN students' educational placement in special education schools and regular schools and their achievement motivation. SEN students in special education schools reported higher academic self-concepts in German and mathematics as well as more enjoyment of learning than SEN students in regular schools. Placement differences could be explained by the lower class-average achievement in special education schools compared to regular schools. These results are in accordance with the big-fish-little-pond effect (Marsh, 1987), which states that students use social comparisons with their classmates for the formation of their academic self-concept and their enjoyment of learning. Furthermore, SEN students who felt emotionally supported by their classmates reported higher achievement motivation.

The key results of this dissertation point to opposite relations between the educational placement of SEN students with their school achievement and achievement motivation. These diverging results present a challenge for creating an optimal learning environment for students with SEN.

Liste der Beiträge

Erster Einzelbeitrag:

Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A. & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 212–234). Weinheim: Beltz Juventa.

Zweiter Einzelbeitrag:

Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191. doi: 10.1007/s11577-014-0253-x

Dritter Einzelbeitrag:

Kocaj, A., Kuhl, P., Rjosk, C., Jansen, M., Pant, H. A. & Stanat, P. (2015). Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf. In: P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 335–370). Wiesbaden: Springer VS.

Vierter Einzelbeitrag:

Kocaj, A., Kuhl, P., Jansen, M., Pant, H. A. & Stanat, P. (2017). *Educational placement and achievement motivation of students with special educational needs*. Manuskript eingereicht zur Publikation.

1

Einleitung

1. Einleitung

Die schulische Entwicklung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) ist in den letzten Jahren nicht zuletzt durch Ratifizierung der UN-Behindertenrechtskonvention (UN-BRK, 2008) verstärkt in den Blickpunkt der empirischen Bildungsforschung gerückt. In der UN-BRK (2008) setzen sich die Vertragsstaaten zum Ziel, ihre Bildungssysteme inklusiver zu gestalten. In Anlehnung an Artikel 24 der UN-BRK (2008) betont die Kultusministerkonferenz (KMK, 2010) den Ausbau des gemeinsamen Lernens von Kindern und Jugendlichen mit und ohne sonderpädagogischen Förderbedarf (SPF) als eines der zentralen Ziele inklusiver Gestaltungsmaßnahmen. Dieser Ausbau gemeinsamen Lernens wird als eine der umfassendsten Schulreformen im deutschen Bildungssystem in den letzten 20 Jahren erachtet (Baumert, 2016a; Grosche, 2015; Weishaupt, 2016).

Zum einen bedeutet die Umsetzung der UN-BRK (2008) eine Veränderung des gesamten Schulsystems und insbesondere des differenzierten sonderpädagogischen Fördersystems, in dem primär Förderschulen für Kinder und Jugendliche mit SPF zuständig sind (Baumert, 2016a; Powell & Pfahl, 2012). Deutschland besitzt im Vergleich zu anderen europäischen Staaten ein stark leistungsdifferenzierendes Bildungssystem (Organisation for Economic Co-operation and Development, OECD, 2013). Schülerinnen und Schüler werden dabei auf Grundlage ihrer schulischen Leistungen vergleichsweise früh in ihrer Schullaufbahn unterschiedlichen Bildungsgängen zugewiesen (Chmielewski, Dumont & Trautwein, 2013; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006).

Zum anderen gibt die Umsetzung der UN-BRK (2008) auch Impulse zur Veränderung pädagogischer Unterrichtskonzepte (Baumert, 2016a). Leistungsbezogene Differenzierungsmaßnahmen im deutschen Bildungssystem basieren maßgeblich auf der Annahme, dass Schülerinnen und Schüler vom Unterricht in leistungshomogenen Lerngruppen profitieren (Gröhlich, Scharenberg & Bos, 2009; Powell, 2009; Scharenberg, 2012). Im Zuge inklusiver Unterrichtsbemühungen soll dagegen der gemeinsame Unterricht von Schülerinnen und Schülern unterschiedlicher Leistungsniveaus ausgebaut werden, was die Lehrkräfte vor Herausforderungen im Umgang mit Leistungsheterogenität und damit verbundenen Methoden der Binnendifferenzierung stellt (Tillmann & Wischer, 2006). Ein Ziel von schulischer Inklusion ist es, schulische Organisationsformen und pädagogische Praktiken so zu verändern, dass die Heterogenität der Schülerschaft wertgeschätzt und angemessen berücksichtigt wird

(Ainscow, Booth & Dyson, 2006; Hinz, 2010). Gleichzeitig sollen alle Schülerinnen und Schüler einer Lerngruppe optimal gefördert werden (Grosche, 2015).

Möchte man die Auswirkungen der inklusiven Beschulung empirisch untersuchen, muss zunächst geklärt werden, was darunter verstanden wird. Inklusion hat sich zu einem prägenden Konzept in der deutschen Bildungslandschaft entwickelt, über dessen Umsetzung kontrovers diskutiert wird (Ahrbeck, 2014; Autorengruppe Bildungsberichterstattung, 2014; Hinz, 2009; Werning, 2014). Das liegt auch daran, dass unter dem Begriff Inklusion Unterschiedliches verstanden wird. Je nach Definition umfasst der Begriff Inklusion verschiedene Aspekte; eine einheitlich und präzise Begriffsbestimmung liegt bisher nicht vor (Artiles, Kozleski, Dorn & Christensen, 2006; Werning, 2010). Ein wichtiger Aspekt von schulischer Inklusion ist die gemeinsame Beschulung von Kindern und Jugendlichen mit SPF in allgemeinen Schulen anstatt einer separaten Beschulung in Förderschulen (Hinz, 2009; Werning, 2010; Wild et al., 2015). Im Fokus dieser Dissertation stehen deshalb Zusammenhänge zwischen der Beschulung in Förderschulen oder allgemeinen Schulen und Merkmalen der schulischen Entwicklung von Schülerinnen und Schülern mit SPF.

Neben der begrifflichen Definition von Inklusion muss weiterhin konkretisiert werden, auf welche Schülergruppe sich die empirische Betrachtung von inklusiver Beschulung bezieht. In dieser Dissertation werden Schülerinnen und Schüler mit SPF betrachtet. Kinder und Jugendliche mit SPF bilden eine heterogene Schülergruppe. Ein sonderpädagogischer Förderbedarf wird diagnostiziert, wenn Schülerinnen und Schüler in ihren Entwicklungs- und Lernmöglichkeiten erheblich beeinträchtigt sind und angenommen wird, dass sie ohne besondere sonderpädagogische Unterstützung nicht erfolgreich am Regelunterricht der allgemeinen Schule teilnehmen können (KMK, 1994). Dabei werden bundesweit bis zu acht verschiedene Förderschwerpunkte unterschieden (Autorengruppe Bildungsberichterstattung, 2014). Im Fokus der vorliegenden Dissertation stehen Schülerinnen und Schüler mit den Förderschwerpunkten *Lernen*, *Sprache* und *emotionale und soziale Entwicklung*. Insbesondere für diese drei Förderschwerpunkte wird zunehmend eine Abschaffung von Förderschulen diskutiert und gefordert, Kinder und Jugendliche mit diesen Förderschwerpunkten stattdessen in allgemeinen Schulen zu unterrichten (Autorengruppe Bildungsberichterstattung, 2014; Gresch, Piezunka & Solga, 2014). Entsprechend stieg der Anteil an Schülerinnen und Schülern mit SPF in allgemeinen Schulen in den letzten 15 Jahren kontinuierlich an (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2015).

Die zunehmende Beschulung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen führt dazu, dass diese Schülergruppe verstärkt in Schulleistungsstudien einbezogen werden (Heydrich, Weinert, Nusser, Artelt & Carstensen, 2013; Stanat et al., 2012; Werning, 2014). Grundlage dieser Dissertation ist der IQB-Ländervergleich 2011 (Stanat et al., 2012), ein bundesweites Large-Scale Assessment in der Primarstufe. Anhand der Daten des IQB-Ländervergleichs 2011 werden Zusammenhänge zwischen Beschulungsart und schulischen Outcomes von Heranwachsenden mit SPF untersucht. Um Aussagen über Schülerinnen und Schüler mit SPF treffen zu können, müssen die Test- und Befragungsinstrumente für diese Schülergruppe reliabel und valide messen. Allerdings wurden die Leistungstests und Kontextfragebögen im IQB-Ländervergleich 2011 ursprünglich nicht für die Schülergruppe mit SPF konzipiert. Das wirft die Frage auf, inwieweit die Testinstrumente auch für die Schülergruppe mit SPF geeignet sind und valide Schlussfolgerungen über deren schulische Outcomes ermöglichen.

Die Frage, in welcher Schulart Kinder und Jugendliche mit SPF höhere schulische Leistungen beziehungsweise eine höhere schulische Motivation erzielen, wird kontrovers diskutiert. Auf der einen Seite verweisen Fürsprecherinnen und Fürsprecher einer separaten Beschulung von Schülerinnen und Schülern mit SPF darauf, dass Förderschulen einen Schonraum für Kinder und Jugendliche mit SPF darstellen (Schumann, 2007). Lehrkräfte in Förderschulen erhalten eine spezialisierte Ausbildung, um auf die besonderen Bedürfnisse der Schülergruppe mit SPF einzugehen (Ahrbeck, 2014). Außerdem sind die Lehrpläne und Unterrichtsmaterialien in Förderschulen stärker auf die Lernvoraussetzungen von Kindern und Jugendlichen mit SPF ausgerichtet (Vaughn, Elbaum & Schumm, 1996). Darüber hinaus lernen Schülerinnen und Schüler mit SPF in Förderschulen gemeinsam mit Schülerinnen und Schülern mit ähnlichen Lernvoraussetzungen. Dadurch könnten sie eher schulische Erfolge erleben und werden vor zu hohen Leistungsanforderungen und Konkurrenzdruck geschützt (Schumann, 2007). Von dieser geschützten Lernumgebung in Förderschulen könnten Schülerinnen und Schüler mit SPF in ihrer schulischen Entwicklung profitieren.

Auf der anderen Seite verweisen Befürworterinnen und Befürworter einer inklusiven Beschulung, dass Kinder und Jugendliche mit SPF in allgemeinen Schulen eine anregende Lernumgebung vorfinden, die sich positiv auf ihre schulische Entwicklung auswirken könnte (Markussen, 2004, Wocken, 2005). Die Lernumgebung in allgemeinen Schulen ist im Vergleich zu Förderschulen durch anspruchsvollere Lehrpläne, höhere Erwartungshaltungen der schulischen Akteure und leistungsstärkere Mitschülerinnen und Mitschüler gekennzeich-

net (Peetsma, Vergeer, Roeleveld & Karsten, 2001). Von dieser anregenderen Lernumgebung in allgemeinen Schulen könnten Schülerinnen und Schüler mit SPF profitieren.

Weiterhin erhielten Kinder und Jugendliche mit SPF in allgemeinen Schulen bessere Bildungschancen als in Förderschulen (Powell & Pfahl, 2012). Die Mehrheit der Schülerinnen und Schüler mit SPF in Förderschulen erreicht keinen Schulabschluss, der zu berufsqualifizierenden Abschlüssen der allgemeinen Schule äquivalent ist (Powell & Pfahl, 2012). Dadurch erfolge der Übergang in den Ausbildungs- beziehungsweise Arbeitsmarkt unter erschwerten Bedingungen oder könnte gänzlich misslingen (Pfahl, 2012). Diese typischen Bildungskarrieren von Schülerinnen und Schülern an Förderschulen ließen Zweifel daran aufkommen, dass eine sonderpädagogischen Förderung in Förderschulen für die gesellschaftliche Teilhabe im Erwachsenenalter förderlich ist (Powell & Pfahl, 2012). Eine gemeinsame Beschulung in allgemeinen Schulen ermögliche hingegen Bildungswege, die Schülerinnen und Schülern mit SPF an Förderschulen zum großen Teil verwehrt werden (Pfahl, 2012).

Zusammenfassend können sowohl Argumente für höhere schulische Kompetenzen beziehungsweise eine höhere schulische Motivation von Kindern und Jugendlichen mit SPF in Förderschulen als auch in allgemeinen Schulen angebracht werden. Insgesamt kann noch keine abschließende Einschätzung über den Einfluss der Beschulungsart auf die schulischen Leistungen und die schulische Motivation von Schülerinnen und Schülern mit SPF getroffen werden (Grosche & Vock, im Druck). Darüber hinaus werden in den meisten Studien ausschließlich Kinder und Jugendliche mit dem Förderschwerpunkt *Lernen* betrachtet. Dabei bleibt offen, ob sich diese Ergebnisse auch auf Schülerinnen und Schüler mit den Förderschwerpunkten *Sprache* und *emotionale und soziale Entwicklung* übertragen lassen.

In der vorliegenden Dissertation wird untersucht, wie die Beschulungsart mit Merkmalen der schulischen Entwicklung von Kindern mit SPF in der Primarstufe zusammenhängt. Sie ist wie folgt gegliedert: Im Theorieteil (Kapitel 2) wird nach einem Überblick über das sonderpädagogische Fördersystem in Deutschland (Kapitel 2.1) und einer Beschreibung verschiedener Inklusionsdefinitionen (Kapitel 2.2) die Situation der Schülerinnen und Schüler mit SPF in Deutschland dargestellt (Kapitel 2.3). Dabei wird auf die Definition eines SPF sowie auf Feststellungsverfahren und daran anschließende Schulartzuweisungen eingegangen. Daran schließt sich eine Beschreibung der Schülergruppe mit SPF in Deutschland auf Basis amtlicher Schulstatistiken an, die mit dem Anteil an der gesamten Schülerpopulation und der typischen Lernumgebung von Schülerinnen und Schülern mit SPF in anderen Ländern verglichen wird. Anschließend werden methodische Herausforderungen beim Einbezug der

Schülergruppe mit SPF in Schulleistungsstudien beschrieben und darauf eingegangen, wie diesen Herausforderungen begegnet werden kann (Kapitel 2.4). In Kapitel 2.5 werden auf Grundlage theoretischer Annahmen zu schulartspezifischen Lernumgebungen Argumente für eine separate Beschulung von Kindern und Jugendlichen mit SPF in Förderschulen und für eine gemeinsame Beschulung in allgemeinen Schulen gegenübergestellt. Dabei liegt ein Schwerpunkt auf der differenzierten Betrachtung möglicher Einflüsse der Beschulungsart auf die schulischen Leistungen und die schulische Motivation von Schülerinnen und Schülern mit SPF.

An die Beschreibung theoretischer Annahmen und relevanter Forschungsbefunde knüpft die Beschreibung der Ziele und Fragestellungen der vorliegenden Dissertation an (Kapitel 3). Anschließend werden die vier Teilstudien vorgestellt, aus der sich die vorliegende Dissertation zusammensetzt (Kapitel 4). In der ersten Teilstudie prüften wir, ob die im IQB-Ländervergleich 2011 eingesetzten Kompetenztests geeignet sind, die schulischen Kompetenzen von Kindern mit SPF zu erfassen. Darauf aufbauend untersuchten wir, ob sich die schulischen Kompetenzen (Teilstudie 2 und 3) und die schulische Motivation (Teilstudie 4) von Kindern und Jugendlichen mit SPF in Abhängigkeit von der Beschulungsart unterscheiden. Dabei berücksichtigten wir Unterschiede in der Klassenzusammensetzung zwischen Förderschulen und allgemeine Schulen als mögliche Erklärungen für Schulartunterschiede. In der abschließenden Diskussion (Kapitel 5) werden die zentralen Befunde der Dissertation zusammengefasst, interpretiert und in den Forschungsstand eingeordnet (Kapitel 5.1). Danach wird auf inhaltliche und methodische Herausforderungen des gewählten Untersuchungsansatzes eingegangen (Kapitel 5.2) und es werden die gegenläufigen Zusammenhänge zwischen Beschulungsart und schulischen Kompetenzen sowie schulischer Motivation von Kindern mit SPF diskutiert (Kapitel 5.3). Die Diskussion schließt mit der Beschreibung von Implikationen für zukünftige Forschungsvorhaben (Kapitel 5.4) und für die pädagogische Praxis (Kapitel 5.5) sowie einem Fazit (Kapitel 5.6).

2

Theoretischer Hintergrund

2. Theoretischer Hintergrund

2.1 Historischer Überblick über das sonderpädagogische Fördersystem in Deutschland

Deutschland besitzt ein stark ausgebautes und differenziertes sonderpädagogisches Fördersystem (Powell & Pfahl, 2012). Förderschulen gehören zwar zum allgemeinen Bildungswesen, sie bilden aber schulrechtlich ein eigenes Schulsystem mit Verankerung im Bildungs- und im Gesundheitswesen (Pfahl, 2011).

Deutschland war eines der ersten Länder, das Kindern und Jugendlichen mit Behinderungen Zugang zum Bildungssystem ermöglichte (Pfahl, 2011). Zu diesem Zweck wurden Anfang des 20. Jahrhunderts Sonder- beziehungsweise Hilfsschulen gegründet, die die allgemeinen Schulen entlasten und eine optimale Förderung für Kinder und Jugendliche mit Behinderungen ermöglichen sollten (Ellger-Rüttgardt, 2008). Damit verbunden erfolgte der Ausbau des Hilfsschullehramtes zu einem vom allgemeinen Schulwesen separaten Berufsstand (Hänsel, 2005). Der Berufsverband der Sonderpädagoginnen und Sonderpädagogen erhielt dadurch eine gewichtige und unabhängige Stimme (Hänsel, 2005; Pfahl, 2011). Nach dem 2. Weltkrieg wurde das sonderpädagogische Fördersystem ausgebaut und ausdifferenziert (Powell & Pfahl, 2012). Für den Ausbau der Förderschulen und für eine separate Beschulung von Kindern und Jugendlichen mit SPF machten sich insbesondere professionelle Verbände der Sonderpädagoginnen und Sonderpädagogen stark (Schumann, 2007). Die starke Position zugunsten einer separaten Beschulung in Förderschulen dürfte auf pädagogischen Überzeugungen, aber auch auf institutionelle Faktoren, wie beispielsweise die selbsterhaltenden und selbstverstärkenden Motive sonderpädagogischer Verbände, zurückzuführen sein (Pfahl, 2011; Powell, 2009).

Das historisch gewachsene, differenzierte sonderpädagogische Fördersystem steht im Einklang mit dem mehrgliedrigen, differenzierten Schulsystem in Deutschland. Das deutsche Bildungssystem zeichnet sich im Vergleich zu den Bildungssystemen anderer europäischer Staaten durch eine starke und vielfältige Leistungsdifferenzierung aus (OECD, 2013). Schülerinnen und Schüler werden in Deutschland schon sehr früh in ihrer Schullaufbahn auf unterschiedliche Schularten mit eigenen Lehrplänen und Stundentafeln verteilt (OECD, 2013). Dazu zählt auch die Überweisung von Kindern und Jugendlichen auf Förderschulen beziehungsweise separate Förderklassen. Der Ausbau eines differenzierten Förderschulsys-

tems in Deutschland hatte zum Ziel, die heterogenen Bedürfnisse der Schülergruppen mit Behinderungen gerecht zu werden und ihnen eine entsprechende Schulbildung zu ermöglichen. Dieses Ziel sollte durch die Bildung leistungshomogener Lerngruppen in verschiedenen Bildungsgängen erreicht werden (Opp, Budnik & Fingerle, 2008; Powell, 2009). Dabei wird argumentiert, dass der Unterricht in leistungshomogenen Lerngruppen besser an die Lernvoraussetzungen und Bedürfnisse von Kindern und Jugendlichen angepasst werden kann (Gröhlich et al., 2009). Die empirische Evidenz für die Vorteile leistungshomogener Gruppen ist allerdings nicht eindeutig (Gröhlich et al., 2009; Köller, Schütte, Zimmermann, Retelsdorf & Leucht, 2013, Tillmann & Wischer, 2006). Die starken Differenzierungsbemühungen im deutschen Bildungssystem scheinen vor allem auch in tief verwurzelten pädagogischen Konzepten und Traditionen begründet zu liegen (Bos, Müller & Stubbe, 2010; Powell, 2009; Hinz, 2002).

Bemühungen hin zur gemeinsamen Beschulung von Kindern und Jugendlichen mit und ohne SPF gingen in den 1970er Jahren von Elterninitiativen und reformpädagogischen Bewegungen aus und führten zu ersten Pilotprojekten und Modellschulen (Biermann & Powell, 2014; Preuss-Lausitz, 2001). Zudem entwickelten sich regionale Unterschiede in der Umsetzung des gemeinsamen Unterrichts, die sich noch heute in Unterschieden zwischen den Bundesländern im Anteil an Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen widerspiegeln (Autorengruppe Bildungsberichterstattung, 2016). Zusätzlichen Auftrieb erfuhren Bemühungen eines gemeinsamen Unterrichts durch internationale Konventionen wie die Salamanca-Erklärung (United Nations Educational, Scientific and Cultural Organization, UNESCO, 1994), die Charta von Luxemburg (Europäische Kommission, 1996) und die UN-BRK (2008, s. Kapitel 2.2.1), die von der KMK in ihren bundesweiten Empfehlungen zur sonderpädagogischen Förderung aufgegriffen wurden (z. B. KMK, 1994, 2010, 2011). In diesen Konventionen wurden auch die Begriffe einer integrativen beziehungsweise inklusiven Beschulung eingeführt, die in der Bildungsforschung und -praxis jedoch unterschiedlich definiert werden.

2.2 Inklusionsdefinitionen

Es gibt verschiedene Möglichkeiten, sich dem Begriff Inklusion anzunähern (Artiles et al., 2006; Werning, 2010); eine einheitliche Definition liegt bisher nicht vor (Dyson, Howes & Roberts, 2002; Göransson & Nilhom, 2014a; Grosche, 2015). Stattdessen werden häufig zentrale Merkmale beziehungsweise Aspekte von Inklusion angeführt. So beschreiben Ainsworth et al. (2005) als zentrale Merkmale von Inklusion:

- a. die steigende Teilhabe von marginalisierten Gruppen und die Verringerung exklusiver Praktiken in allen gesellschaftlichen Bereichen
- b. Veränderungen schulischer Organisationsformen und pädagogischer Praktiken, die dazu führen, dass die Heterogenität der Schülerschaft angemessen berücksichtigt und wertgeschätzt wird (s. a. Hinz, 2010; Prengel, 2013)

Im deutschsprachigen Raum wird unter Inklusion vor allem die Beschulung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen anstatt in Förderschulen verstanden (Grosche, 2015; Werning, 2014). In diesem Kapitel wird zunächst die UN-Behindertenrechtskonvention vorgestellt, die die Definitionen von Inklusion in der Bildungsforschung maßgeblich beeinflusst hat. Darauf aufbauend werden verschiedene Definitionen von Inklusion dargestellt und systematisiert. Daran schließt sich eine Abgrenzung der Begriffe Inklusion und Integration an. Abschließend wird die der vorliegenden Dissertation zugrundeliegende Definition von Inklusion unter Berücksichtigung der verschiedenen Definitionen diskutiert.

2.2.1 Definition von Inklusion in der UN-Behindertenrechtskonvention

Die UN-Behindertenrechtskonvention (UN-BRK, 2008) ist eine zentrale Referenz für Definitionen von Inklusion (Hopf & Kronauer, 2016; Wrase, 2015). Mit der Ratifizierung der UN-BRK (2008) verpflichten sich alle Vertragsstaaten zur Umsetzung eines inklusiven Bildungssystems (Art. 24, Absatz 1: „*States Parties shall ensure an inclusive education system*“). Die UN-BRK wurde in Deutschland per Gesetz vom 21.12.2008 ratifiziert und erlangte am 26.03.2009 den Rang eines Bundesgesetzes. Das Hauptziel der UN-BRK (2008) ist der Abbau von exklusiven Praktiken in allen Gesellschaftsbereichen und die Stärkung des Zugangs und der Teilhabe von Menschen mit Behinderungen (Wrase, 2015). Die Gestaltung

inklusive Bildungssysteme wird in Artikel 24 (Absatz 2) geregelt und anhand der folgenden Merkmale beschrieben:

- a. Verzicht auf Ausschluss und Diskriminierung von Menschen mit Behinderungen in allen Bereichen des allgemeinen Bildungssystems (Primar- und Sekundarstufe, Hochschulbildung, etc.)
- b. Umsetzung eines wohnortnahen Unterrichts für Menschen mit Behinderungen gemeinsam mit Menschen ohne Behinderung
- c. Gestaltung der Lernumgebung innerhalb des allgemeinen Bildungssystems, so dass individuelle Bedürfnisse von Menschen mit Behinderungen angemessen berücksichtigt werden (z. B. durch den Abbau von Barrieren)
- d. Gewährleistung individuell angepasster Unterstützung innerhalb des allgemeinen Bildungssystems, so dass Menschen mit Behinderungen in ihrer akademischen und sozialen Entwicklung profitieren.

Der Fokus in Artikel 24 der UN-BRK (2008) liegt auf Menschen mit Behinderungen und der Schulart, in der sie unterrichtet werden. Latham und Watkins (2014) interpretieren die Forderung einer inklusiven Schulbildung als Umsetzung des Rechts auf den Besuch einer wohnortnahen Schule ohne vorherige Selektion. Auch Wrase (2015) stellt fest, dass die zentrale Implikation der UN-BRK (2008) für die schulische Organisation darin bestehe, einen gemeinsamen Unterricht von Schülerinnen und Schülern mit und ohne SPF zu ermöglichen. Dabei sollten für Schülerinnen und Schüler mit SPF individuelle Unterstützungsangebote bereitgestellt werden, die ihnen eine erfolgreiche Teilnahme am Unterricht ermöglichen (vgl. Art. 24, UN-BRK, 2008).

Es wird jedoch kontrovers diskutiert, ob die UN-BRK (2008) eine komplette Abschaffung der Förderschulen vorsieht. Wrase (2016) interpretiert die UN-BRK (2008) als Aufruf zur kompletten Auflösung der Förderschulen und verweist darauf, dass jegliche Form einer separaten Beschulung eine Diskriminierung aufgrund einer Behinderung darstelle (s. a. UN Committee on the Rights of Persons with Disabilities, 2015). Im Gegensatz dazu weisen Powell und Pfafl (2012) darauf hin, dass Förderschulquoten von 5 bis 20 Prozent in Einklang mit der UN-BRK (2008) stehen könnten (Poscher, Rux & Langer, 2008). Auch Fischer, Heimlich, Kahlert und Lelgemann (2014) argumentieren in einer wissenschaftlichen Expertise im Auftrag des bayerischen Landtags, dass die Aufrechterhaltung von Förderschulen nicht gegen die UN-BRK (2008) verstoße. Begründet wird diese Ansicht damit, dass eine separate Beschulung keine Diskriminierung darstelle, wenn sie durch effektive Förderung zur Herstel-

lung von Gleichberechtigung beiträgt (Fischer et al., 2014). Demnach sei die Förderung von Kindern und Jugendlichen mit SPF nicht an einen bestimmten Förderort geknüpft. Im Gegenteil würde die Bereitstellung einer Vielfalt von Förderorten und die damit verbundene Wahlfreiheit eine optimale Förderung ermöglichen und die Selbstbestimmung von Menschen mit Behinderungen stärken (Fischer et al., 2014). Konsens scheint zumindest dahingehend zu bestehen, dass die Beschulung von Kindern und Jugendlichen mit SPF in Förderschulen eine Ausnahme darstellen sollte (s. a. Salamanca-Erklärung, UNESCO, 1994) und nur dann vertretbar ist, wenn ein gemeinsamer Unterricht nur mit unverhältnismäßig großem Aufwand erreicht werden könnte oder im Einzelfall eine inklusive Beschulung unzumutbar wäre (Wrase, 2015).

2.2.2 Definitionen von Inklusion in der empirischen Bildungsforschung

In der empirischen Bildungsforschung und in verwandten Disziplinen existieren verschiedene Definitionen von Inklusion. In Anlehnung an die UN-BRK (2008) beziehen sich viele Definitionen von Inklusion auf den Förderort von Schülerinnen und Schülern mit SPF. Göransson und Nilholm (2014a) untersuchten, wie Inklusion in aktuellen empirischen Studien definiert wird. Sie unterscheiden vier Gruppen von Inklusionsdefinitionen. Die erste Gruppe von Definitionen betrachtet *Inklusion als Platzierung*, also als die Teilnahme von Schülerinnen und Schülern mit SPF im gemeinsamen Unterricht der allgemeinen Schule (s. a. Florian, 2014). Diese Definition findet sich auch in einer einflussreichen Zusammenfassung des Forschungsstandes zu Effekten von Inklusion wieder (Ruijs & Peetsma, 2009, S. 67):

“Inclusive education can be defined as educating children with special educational needs in regular schools, instead of in special schools.”

Unterschiede zwischen verschiedenen Definitionen dieser Gruppe bestehen darin, wie viel Zeit Schülerinnen und Schüler mit SPF im gemeinsamen Unterricht verbringen (z. B. über 80% des Schultages im gemeinsamen Unterrichts; Florian, 2014). Neben der Platzierung in allgemeinen Schulen werden in dieser Gruppe von Inklusionsdefinitionen keine weiteren Merkmale der Lernumgebung berücksichtigt. Die meisten Forscherinnen und Forscher stimmen jedoch darüber überein, dass Inklusion mehr als die Beschulung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen bedeuten sollte (Florian, 2014; Werning, 2010).

Die zweite Gruppe von Definitionen beschreibt *Inklusion als Eingehen auf soziale und akademische Bedürfnisse von Schülerinnen und Schülern mit SPF*. Inklusion wird nicht nur als Platzierung von Kindern und Jugendlichen mit SPF in allgemeinen Schulen beschrieben, sondern darüber hinaus wird auch die Ausgestaltung der sonderpädagogischen Förderung (z. B. Lehrpläne, Unterstützungsressourcen) in den Blick genommen (z. B. Mitchell, 2013). Außerdem sind die schulischen Outcomes von Schülerinnen und Schülern mit SPF zentraler Bestandteil der Definition von Inklusion. Verschiedene Definitionen in dieser Gruppe unterscheiden sich dahingehend, was unter der Berücksichtigung sozialer und kognitiver Bedürfnisse verstanden wird (z. B. Erreichen von Bildungsstandards, Erreichen individueller Lernziele, Wohlbefinden, Beziehungen zu Mitschülerinnen und Mitschülern, soziale Kompetenzen etc.). Eine Schwäche dieser Definitionen besteht darin, dass Merkmale der Lernumgebung mit Zielen inklusiver Bemühungen vermischt werden. Das erschwert die Evaluation inklusiver Bemühungen in empirischen Studien, da positive schulische Outcomes von Schülerinnen und Schülern mit SPF nicht das Ergebnis von Inklusion, sondern einen Teil der Definition von Inklusion bilden. Inklusionsdefinitionen dieser Gruppe beziehen sich so wie die erste Gruppe von Definitionen ausschließlich auf Schülerinnen und Schüler mit SPF.

Die dritte Gruppe von Definitionen versteht *Inklusion als Berücksichtigung sozialer und akademischer Bedürfnisse der gesamten Schülerschaft*. Inklusion bezieht sich also nicht nur auf Schülerinnen und Schüler mit SPF, sondern umfasst alle Schülerinnen und Schüler in heterogenen Lerngruppen. Vertreterinnen und Vertreter solcher Definitionen betonen, dass Inklusion sich nicht nur auf sonderpädagogische Förderung bezieht, sondern generell Prinzipien zur Schulentwicklung umfasst, also Hinweise zur Änderung von Schulpraktiken dahingehend gebe, dass allgemeine Schulen eine größere Vielfalt von Schülerinnen und Schülern unterrichten können (Ainscow, 1991; Florian, 2014). In Einklang mit der Definition von Biewer (2010) und Hinz (2010) beziehen sich Definitionen dieser Gruppe nicht nur auf Kinder und Jugendliche mit SPF. Vielmehr sollen bei der Unterrichtsgestaltung weitere relevante Dimensionen von Heterogenität (z. B. Zuwanderungshintergrund, sozialer Hintergrund, Leistungsstärke) berücksichtigt werden. Alle Schülerinnen und Schüler werden dabei gemeinsam unterrichtet und erhalten bei Bedarf zusätzliche Unterstützung von Sonderpädagoginnen und Sonderpädagogen oder auch allgemeinen Pädagoginnen und Pädagogen (Artiles et al., 2006).

Die vierte Gruppe von Definitionen beschreibt *Inklusion anhand von Merkmalen auf gesellschaftlicher Ebene*. Inklusion bezieht sich dabei nicht nur auf den schulischen Kontext,

sondern beinhaltet die Wertschätzung von Vielfalt in allen Lebensbereichen der Gesellschaft (Grosche & Vock, im Druck). Diese Lebensbereiche umfassen neben der Schule unter anderem die Nachbarschaft und die Gemeindeebene, wobei Akteure unterschiedlicher Bereiche miteinander kooperieren (Naraian, 2011). Ein weiterer Aspekt dieser Inklusionsdefinitionen ist die Verbesserung der sozialen Teilhabe aller Menschen, die von Marginalisierung betroffen sind. Dieser Aspekt entspricht einer prozessorientierten Sicht auf Inklusion, bei der strukturelle Änderungen im Bildungssystem und in weiteren Lebensbereichen im Fokus stehen anstatt Inklusion als einen abgeschlossenen Zustand zu definieren. Inklusionsdefinitionen in dieser Gruppe unterscheiden sich entsprechend dem großen Anwendungsbereich dahingehend, welche Aspekte zur Beurteilung einer inklusiven Gesellschaft herangezogen werden (z. B. Gleichberechtigung, Gerechtigkeit, Wertschätzung von Verschiedenheit) und welche Ebene (z. B. Nachbarschaft, Gemeinde, Gesellschaft) in den Blick genommen wird. Definitionen der vierten Gruppe entsprechen dem umfassendsten Verständnis von Inklusion. Sie lassen sich in empirischen Studien aber nur schwer operationalisieren, da sie sehr global formuliert sind und hohe Anforderungen an die Datenerhebung stellen. So wären ein Mehrebenenansatz (z. B. Gemeinschaftsebene, Schulebene, Klassenebene, Individualebene) und ein längsschnittliches Studiendesign erforderlich, um den Prozess hin zu einer inklusiven Gesellschaft abzubilden (European Agency for Development in Special Needs Education, 2011; Tomlinson, 2012).

2.2.3 Abgrenzung der Begriffe Inklusion und Integration

Neben der Vielfalt an Definitionen von Inklusion besteht eine weitere Herausforderung darin, die Begriffe Inklusion und Integration voneinander abzugrenzen. Beide Begriffe werden häufig synonym verwendet und ähnelten sich in ihrer ursprünglichen Verwendung (Grosche, 2015; Wocken, 2009). Bereits in den 1970er Jahren etablierte sich der Begriff Integration in Deutschland (Powell & Pfahl, 2012). Zu dieser Zeit waren die Ideen der Integrationsbewegung nahezu identisch mit denen der Inklusion (Hinz, 2004). So bezog sich auch Integration auf die Beachtung und Wertschätzung sämtlicher Heterogenitätsdimensionen (Pfahl, 2011). Der gemeinsame Unterricht von Kindern und Jugendlichen mit SPF in allgemeinen Schulen wurde lediglich als erster Schritt zur Erreichung dieser Ziele gesehen (Grosche, 2015). Im Laufe der Zeit wurde jedoch unter Integration primär die gemeinsame Beschulung von Schülerinnen und Schülern mit SPF verstanden. In Anlehnung an die internationale Debatte

um die Stärkung der Rechte von Menschen mit Behinderungen etablierte sich der Begriff Inklusion im deutschsprachigen Raum (Grosche, 2015). Jedoch wurde der Begriff *Inclusion* in internationalen Erklärungen wie der Salamanca-Erklärung (UNESCO, 1994), der Charta von Luxemburg (Europäische Kommission, 1996) und der UN-BRK (2008) mit dem Begriff Integration übersetzt. Das trug zur begrifflichen Unklarheit beider Begriffe im deutschsprachigen Raum bei (Hinz, 2004).

Hinz (2002) versteht unter Inklusion und Integration zwei verschiedene pädagogische Konzepte und differenziert sie anhand ihrer Umsetzung in der schulischen Praxis (s. Tabelle 1). Im Schulalltag nimmt Integration demnach eine individuumbezogene Sichtweise ein: Schülerinnen und Schüler mit SPF werden mit Hilfe sonderpädagogischer Unterstützung in den Schulalltag der allgemeinen Schule einbezogen. Die Zuweisung sonderpädagogischer Unterstützung ist dabei an die individuelle Feststellung eines SPF geknüpft. Der Fokus von Integration liegt auf einer Anpassung der Lernumgebung an die speziellen Bedürfnisse von Schülerinnen und Schüler mit SPF, die in allgemeinen Schulen unterrichtet werden. In Schulklassen ohne Schülerinnen und Schüler mit SPF sind keine Veränderungen in der Unterrichtsgestaltung vorgesehen. Beispiele für Integration in der schulischen Praxis sind nach Hinz unter anderem zieldifferente, individualisierte Lehrpläne, die ausschließlich für Kinder und Jugendliche mit SPF gelten, oder die Einrichtung spezieller Integrationsklassen in allgemeinen Schulen.

Dem Konzept der Integration liegt implizit eine Zwei-Gruppen-Theorie zugrunde, die eine Differenzierung zwischen Schülerinnen und Schülern mit SPF und Schülerinnen und Schülern ohne SPF postuliert. Dies hat neben differenzierenden Lehrplänen auch unterschiedliche Zuständigkeiten des pädagogischen Personals zur Folge. Vertreterinnen und Vertreter der allgemeinen Pädagogik arbeiten zwar mit Sonderpädagoginnen und Sonderpädagogen zusammen, jedoch fallen Kinder und Jugendliche mit SPF in den Zuständigkeitsbereich der Sonderpädagogik. Hinz kritisiert an dieser Konzeptualisierung von Integration den Fokus auf Limitationen der Schülerinnen und Schüler mit SPF. So könnte die Kategorisierung von Schülerinnen und Schülern mit SPF zu einer Stigmatisierung und geringeren Anspruchshaltung der Lehrkräfte führen. Außerdem würden integrative Praktiken nach Hinz den Förderort von Schülerinnen und Schülern mit SPF betonen anstatt konkrete Merkmale der Unterrichtsgestaltung zu beschreiben.

Im Gegensatz zur Integration umfasst das Konzept der Inklusion eine systemische Perspektive auf die Unterrichtsgestaltung (s. Tabelle 1 nach Hinz, 2002). Der Unterricht wird so gestaltet,

dass Schülerinnen und Schüler mit unterschiedlichen Lernvoraussetzungen gemeinsam in einer Schule lernen können. Heterogene Lerngruppen sind der Ausgangspunkt für die Unterrichtsgestaltung, unabhängig davon, ob in einer Klasse auch Schülerinnen und Schüler mit SPF unterrichtet werden. Zieldifferente Lehrpläne für einzelne Schülergruppen werden als eine Form der Kategorisierung beziehungsweise Etikettierung abgelehnt. Vielmehr wird ein gemeinsames Curriculum präferiert, das sich an der Heterogenität aller Schülerinnen und Schüler orientiert. Dabei können die Lehrpläne in Teilbereichen durch binnendifferenzierende Methoden individualisiert werden (Prengel, 2013). Anstelle einer Kategorisierung verschiedener Schülergruppen wird die gesamte Schülerschaft in inklusiven Schulen als eine nicht unterteilbare heterogene Lerngruppe betrachtet und auf jegliche Kategorisierung verzichtet. Die Annahme einer unteilbaren Lerngruppe führt zu einer Verschränkung von Sonderpädagogik und allgemeiner Pädagogik, so dass beide Fachbereiche für alle Schülerinnen und Schüler zuständig sind. Im Gegensatz zur individuellen Ressourcenzuweisung stehen sonderpädagogische Ressourcen der gesamten Lerngruppe zur Verfügung. Hinz (2002) sieht einen Vorteil der so verstandenen Inklusion darin, dass ein Paradigmenwechsel auf Schul- und Unterrichtsebene initiiert wird, während Integration lediglich die Zuständigkeit der Sonderpädagogik auf die allgemeine Schule erweitert.

Tabelle 1. *Vergleich zwischen inklusiven und integrativen Praktiken nach Hinz (2002, S. 11-12)*

Integration	Inklusion
individuumszentrierter Ansatz	systemischer Ansatz
Einbezug von Schülerinnen und Schülern mit SPF in allgemeine Schulen	gemeinsame Schule für alle Schülerinnen und Schüler mit heterogenen Ausgangslagen
Anpassungsleistung der Schülerinnen und Schüler mit SPF	Anpassungsleistung der allgemeinen Schule
Zwei-Gruppen-Theorie	Theorie einer pädagogisch nicht unterteilbaren heterogenen Gruppe
Ressourcenzuweisung auf Individualebene	Ressourcenzuweisung auf Schul- bzw. Lerngruppenebene
individualisierte Lehrpläne für Kinder mit SPF	ein gemeinsames Curriculum für die gesamte Schülerschaft mit individualisierten Teilbereichen
Zusammenarbeit zwischen allgemeiner Pädagogik und Sonderpädagogik mit getrennten Zielgruppen	Verschränkung zwischen allgemeiner Pädagogik und Sonderpädagogik mit Zuständigkeit für alle Schülerinnen und Schüler

Inklusion und Integration lassen sich auch anhand eines mehrstufigen Modells zum Umgang mit Menschen mit Behinderungen in der Gesellschaft differenzieren (Grosche, 2015; Hinz, 2002; Powell & Pfahl, 2012). Mit zunehmender Stufe steigt die Anerkennung und Gewährleistung von Menschenrechten für Menschen mit Behinderungen. Während Menschen mit Behinderungen in den ersten beiden Stufen, *Extinktion* beziehungsweise *Exklusion*, kein Recht auf Leben beziehungsweise allgemeine Bildung haben, wird in der dritten Stufe, *Separation/Segregation*, das Recht auf Bildung von Menschen mit Behinderungen durch umfangreiche pädagogische Unterstützung gewährleistet. Allerdings erfolgt die Umsetzung in getrennten Förderschulen (Segregation) oder getrennten Lerngruppen innerhalb einer gemeinsamen Schule (Separation). In der vierten Stufe, der *Integration*, erhalten Menschen mit Behinderungen zusätzlich das Recht auf soziale Teilhabe durch eine gemeinsame Unterrichtung in allgemeinen Schulen, aber die gemeinsame Beschulung muss im Einzelfall begründet werden. In der fünften und höchsten Stufe, der *Inklusion*, wird auf eine Kategorisierung verzichtet und die Aufnahme von allen Schülerinnen und Schülern unabhängig von ihren sozialen Merkmalen und Lernausgangslagen in die allgemeine Schule umgesetzt.

Das Stufenmodell geht von einer Kontinuität zwischen Integration und Inklusion aus. Grosche (2015) führt in seiner Diskussion des Modells aus, dass die fünfte Stufe insbesondere bei kompletter Ablehnung einer getrennten Beschulung und bei vollständigem Verzicht von Feststellungsverfahren eher einem Bildungsideal entspricht. Zudem ist eine Abgrenzung der beiden Stufen Integration und Inklusion aufgrund von Gemeinsamkeiten in der Umsetzung des Rechts auf Bildung in einer gemeinsamen Schule und der Stärkung sozialer Teilhabe aller Schülerinnen und Schüler nicht immer möglich. Deshalb schlägt Grosche (2015) vor, dass die Abgrenzung von Integration und Inklusion in der empirischen Bildungsforschung aufgegeben und konsequent der Begriff Inklusion genutzt werden sollte. Dieses Vorgehen würde auch die historisch begründeten Gemeinsamkeiten beider pädagogischer Konzepte widerspiegeln (Wocken, 2009). Außerdem wäre eine solche Begriffswahl an die internationale Forschungslandschaft anschlussfähiger, in der überwiegend der Begriff Inklusion verwendet wird (Grosche, 2015).

2.2.4 Inklusionsdefinition der vorliegenden Dissertation

In der vorliegenden Dissertation wird auf die Verwendung des Inklusionsbegriffs weitgehend verzichtet. Stattdessen wird von Zusammenhängen zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit SPF gesprochen. Dennoch ist eine Auseinandersetzung mit dem Inklusionsbegriff erforderlich, da die vorliegende Dissertation im Kontext aktueller Inklusionsbemühungen rezipiert wird (Grosche, 2015; Werning, 2014). In dieser Dissertation wird Inklusion als Platzierung von Schülerinnen und Schülern mit SPF in den gemeinsamen Unterricht der allgemeinen Schule definiert. Diese Definition zählt zu den am weitesten verbreiteten Inklusionsdefinitionen in der empirischen Bildungsforschung (Ruijs & Peetsma, 2009). In der Systematisierung von Göransson und Nilholm (2014a) gehört sie zur ersten Gruppe von Inklusionsdefinitionen (vgl. Kap. 2.2.2).

Ein Nachteil dieser Definition besteht darin, dass ausschließlich Schülerinnen und Schüler mit SPF (in den Förderschwerpunkten *Lernen*, *Sprache* und *emotionale und soziale Entwicklung*) betrachtet werden, während weite Inklusionsdefinitionen zusätzliche Merkmale heterogener Lerngruppen berücksichtigen und sich auf die gesamte Schülerschaft einer Schule beziehen. Außerdem lässt sich die vorliegende Definition nicht vom Begriff der Integration abgrenzen, da beide Begriffe die Anwesenheit von Schülerinnen und Schülern mit SPF im gemeinsamen Unterricht der allgemeinen Schule als zentrales definitorisches Merkmal beinhalten (Hinz, 2002). Im Gegensatz dazu beinhalten weite Inklusionsdefinitionen Merkmale der konkreten Ausgestaltung des gemeinsamen Unterrichts (Göransson & Nilholm, 2014a). Dadurch könnten die den Begriffen Inklusion und Integration zugrundeliegenden pädagogischen Konzepte besser differenziert werden (Hinz, 2002). In dieser Dissertation lassen sich Zusammenhänge zwischen der Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit SPF quantifizieren, während die konkrete Ausgestaltung des inklusiven Unterrichts in weiterführenden Studien näher untersucht werden können. Für diese Betrachtung sind längsschnittliche Studien erforderlich, die verschiedene Organisationsformen des gemeinsamen Unterrichts miteinander vergleichen (z. B. Gersten et al., 2009; Wild et al., 2015).

Die hier verwendete Inklusionsdefinition ist aber auch mit Vorteilen für die empirische Untersuchung verbunden. So lässt sich die Platzierung von Kindern und Jugendlichen mit SPF in Förderschulen oder allgemeinen Schulen in großen Schulleistungsstudien eindeutig operationalisieren. Außerdem liegt die hier gewählte Definition von Inklusion der Mehrzahl

empirischer Studien zugrunde. Dadurch können die Ergebnisse der Dissertation leichter in den Forschungsstand eingeordnet werden. Ein weiterer Vorteil besteht darin, dass schulische Outcomes nicht Bestandteil der Inklusionsdefinition sind. Im Vergleich zu weiten Inklusionsdefinitionen (Göransson & Nilholm, 2014a) lassen sich psychosoziale, motivationale und kognitive Outcomes von Schülerinnen und Schülern mit SPF als Effekte von Inklusion betrachten. Im Gegensatz dazu sind schulische Outcomes, wie zum Beispiel hohe schulische Leistungen und eine hohe soziale Integration, Bestandteil weiterer Inklusionsdefinitionen. Merkmale einer inklusiven Lernumgebung lassen sich dadurch nicht von Merkmalen einer gelingenden Inklusion trennen.

2.3 Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

2.3.1 Definition sonderpädagogischer Förderbedarf

Der Begriff des sonderpädagogischen Förderbedarfs (SPF) wird überwiegend im schulischen Kontext verwendet (Autorengruppe Bildungsberichterstattung, 2014). In anderen Lebensbereichen (z. B. vorschulischer Bereich, tertiärer Bildungsbereich, Beruf) werden stattdessen verschiedene Definitionen von Behinderung genutzt. Im Folgenden wird die Definition von Behinderung nach dem Neunten Buch Sozialgesetzbuch und UN-BRK (2008) eingeführt und mit der Definition eines SPF nach Vorgaben der Kultusministerkonferenz (KMK, 1994) verglichen. Im Anschluss werden zentrale Kritikpunkte der KMK-Definition eines SPF benannt. Abschließend werden die in der vorliegenden Arbeit betrachteten Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* begrifflich eingeordnet und von Teilleistungsstörungen abgegrenzt.

2.3.1.1 Definitionen von Behinderung nach dem Neunten Buch Sozialgesetzbuch und der UN-BRK (2008)

Die Feststellung einer Behinderung nach dem Neunten Buch Sozialgesetzbuch (SGB IX, i. d. F. d. Bundesteilhabegesetzes) ist für Menschen mit Behinderungen aus juristischer Perspektive notwendig, um Leistungen in Anspruch zu nehmen und eine Bereitstellung dieser Leistun-

gen einfordern zu können (Autorengruppe Bildungsberichterstattung, 2014). Im Neunten Buch Sozialgesetzbuch wird Behinderung wie folgt definiert (§ 2 Abs. 1 SGB IX):

„Menschen mit Behinderungen sind Menschen, die körperliche, seelische, geistige oder Sinnesbeeinträchtigungen haben, die sie in Wechselwirkung mit einstellungs- und umweltbedingten Barrieren an der gleichberechtigten Teilhabe an der Gesellschaft mit hoher Wahrscheinlichkeit länger als sechs Monate hindern können. Eine Beeinträchtigung nach Satz 1 liegt vor, wenn der Körper- und Gesundheitszustand von dem für das Lebensalter typischen Zustand abweicht. Menschen sind von Behinderung bedroht, wenn eine Beeinträchtigung nach Satz 1 zu erwarten ist.“

Wie bereits dargestellt (Kapitel 2.2.1), hat Artikel 24 der UN-BRK (2008) einen maßgeblichen Einfluss auf die Gestaltung schulischer Lernumgebungen für Schülerinnen und Schüler mit SPF. Dabei wird jedoch an keiner Stelle der Begriff „sonderpädagogischer Förderbedarf“ erwähnt. Stattdessen wird auch in der UN-BRK (2008) der Begriff „Behinderung“ beziehungsweise „Menschen mit Behinderungen“ verwendet. Sowohl die Definition von Behinderung im Neunten Buch Sozialgesetzbuch als auch die Definition von Behinderung in der UN-BRK (2008) folgen dem biopsychosozialen Modell (Köbsell, 2012):

"Zu den Menschen mit Behinderungen zählen Menschen, die langfristige körperliche, seelische oder Sinnesbeeinträchtigungen haben, welche sie in Wechselwirkung mit verschiedenen Barrieren an der vollen, wirksamen und gleichberechtigten Teilhabe an der Gesellschaft hindern können“ (Art. 1, UN-BRK, 2008)

Behinderung ist nach diesem Verständnis das Ergebnis einer Interaktion zwischen Person und Umwelt. Nach dieser Definition können Menschen mit Behinderungen aufgrund der Interaktion ihrer Beeinträchtigungen mit gesellschaftlichen Barrieren nicht in vollem Ausmaß am sozialen Leben teilhaben. Folglich ist das Ziel inklusiver Bemühungen, gesellschaftliche Barrieren abzubauen. Ein Vorteil der biopsychosozialen Sichtweise ist, dass keine künstliche Trennung zwischen einer personenorientierten und gesellschaftlichen Ebene erfolgt (Köbsell, 2012). Außerdem betont das biopsychosoziale Modell, dass Menschen mit Behinderungen das Potenzial haben, am gesellschaftlichen Leben zu partizipieren und es lehnt eine defizitorientierte Sichtweise auf Behinderung ab.

2.3.1.2 Allgemeine Definition eines sonderpädagogischen Förderbedarfs nach KMK (1994)

Sowohl im Neunten Buch Sozialgesetzbuch als auch in der UN-BRK wird der Begriff Behinderung verwendet, während die Empfehlungen der KMK (1994) den Begriff des sonderpädagogischen Förderbedarfs für den schulischen Bereich nutzen. Die KMK (1994) definiert sonderpädagogischen Förderbedarf folgendermaßen:

„Sonderpädagogischer Förderbedarf ist bei Kindern und Jugendlichen anzunehmen, die in ihren Bildungs-, Entwicklungs- und Lernmöglichkeiten so beeinträchtigt sind, daß sie im Unterricht der allgemeinen Schule ohne sonderpädagogische Unterstützung nicht hinreichend gefördert werden können.“ (KMK, 1994; S. 5-6)

Diese Definition stellt zwar individuelle Beeinträchtigungen der Kinder und Jugendlichen in den Vordergrund, verweist aber auch auf die Interaktion individueller Beeinträchtigungen mit der Lernumgebung in allgemeinen Schulen. Die Feststellung einer Behinderung nach dem Neunten Buch Sozialgesetzbuch und die Diagnose eines sonderpädagogischen Förderbedarfs sind voneinander unabhängige Kategorien (Autorengruppe Bildungsberichterstattung, 2014). Wird im Rahmen des Sozialrechts eine Behinderung festgestellt, erfolgt im Anschluss nicht automatisch die Feststellung eines SPF.² Ebenso bestehen bei der Feststellung eines SPF nicht automatisch sozialrechtliche Leistungsansprüche.

In den Empfehlungen der Kultusministerkonferenz zur sonderpädagogischen Förderung von 1994 (KMK, 1994) wurde der bis dahin verwendete Begriff der *Sonderschulbedürftigkeit* durch den Begriff des *sonderpädagogischen Förderbedarfs* ersetzt. Während die Feststellung einer Sonderschulbedürftigkeit eine separate Beschulung in Förderschulen impliziert, soll die Feststellung eines SPF zu einer individualisierten Entscheidung über den Förderort führen. Entsprechend wurde in den Empfehlungen der KMK von 1994 festgehalten, dass Kinder und Jugendliche mit Beeinträchtigungen auch in allgemeinen Schulen unterrichtet werden können, sofern sie dort die benötigten Ressourcen und eine adäquate Förderung erhalten. Bei der Entscheidung über den im Einzelfall angemessenen Förderort empfiehlt die KMK (1994; S. 9), Art und Umfang des Förderbedarfs zu berücksichtigen. Nach Ratifizierung der UN-BRK (2008) aktualisierte die KMK (2010, 2011) ihre Empfehlungen dahingehend, dass die allgemeine Schule für Kinder und Jugendliche mit SPF primär zuständig ist. Der gemeinsame

² In manchen Bundesländern (Hamburg, Bremen, Niedersachsen) ist die Feststellung eines SPF allerdings schulrechtlich an den Behinderungsbegriff nach SGB IX geknüpft (Wrase, 2015).

Unterricht von Kindern und Jugendlichen mit und ohne Behinderung ist demnach zentrales Ziel bildungspolitischer Maßnahmen (KMK, 2011).

2.3.1.3 Definitionen nach Förderschwerpunkt

Die Feststellung eines SPF ist mit der Präzisierung eines Förderschwerpunktes verbunden. In Deutschland werden acht verschiedene Förderschwerpunkte unterschieden: *Lernen, Sprache, emotionale und soziale Entwicklung, geistige Entwicklung, körperliche und motorische Entwicklung, Hören, Sehen* und *chronische Erkrankung* (KMK, 1994, 2011).³ Im Fokus dieser Dissertation stehen die Förderschwerpunkte *Lernen, Sprache* und *emotionale und soziale Entwicklung*, daher wird auch nur auf diese drei Förderschwerpunkte im Folgenden genauer eingegangen.

Der Förderschwerpunkt *Lernen* wird definiert als Beeinträchtigungen im schulischen Lernen und Leistungsverhalten, die durch vielfältige kognitive, sensorische, aber auch soziale und emotionale Ausgangsvoraussetzungen bedingt sein können (KMK, 1999). Es wird angenommen, dass Kinder und Jugendliche mit dem Förderschwerpunkt *Lernen* aufgrund dieser Beeinträchtigungen Ziele und Inhalte der Lehrpläne der allgemeinen Schule im Vergleich zu ihrer Alterskohorte nicht beziehungsweise nur ansatzweise erreichen können. Die Definition des Förderschwerpunkts *Sprache* ist enger umrissen und bezieht sich auf funktionale sprachliche Probleme von Kindern und Jugendlichen, die sich sowohl in der Kommunikation als auch in altersunüblichen Leistungen im schriftlichen und sprachlichen Bereich äußern (KMK, 1998). Der Förderschwerpunkt *emotionale und soziale Entwicklung* wird definiert als Schwierigkeiten im emotionalen Erleben und sozialen Verhalten von Kindern und Jugendlichen, die in familiären und schulischen Interaktionsprozessen hervortreten (KMK, 2000). Ein Förderschwerpunkt *emotionale und soziale Entwicklung* wird bei Schülerinnen und Schülern diagnostiziert, die Probleme in den Bereichen soziale Kompetenz, Erlebnis- und Wahrnehmungsfähigkeit, emotionale Ausdrucksfähigkeit und Selbstregulation haben (KMK, 2000).

³ Zudem wird in manchen Bundesländern ein separater Förderschwerpunkt *autistisches Verhalten* (z. B. Berlin, Nordrhein-Westfalen) ausgewiesen (KMK, 2000). Eine Sonderrolle nimmt auch der Unterricht von chronisch kranken Schülerinnen und Schülern ein, welcher häufig in einem krankenhaushnahen Setting beziehungsweise zu Hause stattfindet (KMK, 1998).

Ein sonderpädagogischer Förderbedarf im Bereich *Lernen* (und zum Teil auch in den Bereichen *Sprache* und *emotionale und soziale Entwicklung*) kann dem Oberbegriff „Lernstörungen“ zugeordnet werden, welcher „*Minderleistungen beim absichtsvollen Lernen*“ (Lauth, Brunstein & Grünke, 2014, S. 17) umfasst. Dabei wird ein SPF von anderen Lernstörungen beziehungsweise Lernschwierigkeiten auf Grundlage der Dimensionen (1) Umfang und (2) zeitliche Dauer abgegrenzt (Klauer & Lauth, 1997; Lauth et al., 2014). Die erste Dimension bezieht sich darauf, ob eine Lernstörung nur bereichsspezifisch (z. B. in einzelnen Schulfächern) oder bereichsübergreifend beziehungsweise allgemein festgestellt wird. Die Diagnose eines SPF im Bereich *Lernen* setzt voraus, dass sich die Beeinträchtigung des schulischen Lernens in verschiedenen Schulfächern beobachten lässt. Im Gegensatz dazu beziehen sich Teilleistungsstörungen (z. B. Lese-Rechtschreibstörung, Rechenstörung) auf bereichsspezifische Domänen. Die zweite Dimension bezieht sich darauf, ob eine Lernstörung vorübergehend oder überdauernd festgestellt wird. Bei einem SPF geht man von einer überdauernden Beeinträchtigung des schulischen Lernens aus. Diese zeitliche Komponente trifft auch auf Teilleistungsstörungen zu, jedoch nicht auf Lernschwierigkeiten, die auf Lernrückstände in einzelnen Fächern zurückzuführen sind (vorübergehend und bereichsspezifisch) oder auf Schulschwierigkeiten aufgrund einer psychischen Belastung im familiären Umfeld (vorübergehend und bereichsübergreifend). Nach Wrase (2015) fallen Schülerinnen und Schüler mit Teilleistungsschwächen beziehungsweise sonstigem Unterstützungsbedarf nicht in den Anwendungsbereich der UN-BRK (2008).

2.3.1.4 Kritik an der KMK-Definition des sonderpädagogischen Förderbedarfs

Die Empfehlungen der KMK (1994) wurden als eine „kopernikanische Wende“ im Umgang mit Kindern und Jugendlichen mit Behinderung bezeichnet (Bleidick, Rath & Schuck, 1995). Zu dieser Einschätzung trug vor allem bei, dass Förderschulen nicht mehr als bevorzugter Förderort für Kinder und Jugendliche mit Behinderungen angesehen wurden. Allerdings wird die Definition eines SPF in den Empfehlungen der KMK (1994) aus verschiedenen Gründen auch kritisiert (Kottmann, 2006; Powell, 2004).

Erstens verweise der Begriff „SPF“ ebenso wie der vorherige Begriff „Sonderschulbedürftigkeit“ auf eine Zuständigkeit der Sonderpädagogik und eine Überlegenheit ihrer Förderangebote (Kottmann, 2006; Powell, 2004). Schülerinnen und Schüler mit einem festgestellten SPF benötigten demnach eine intensivere Förderung, die von der allgemeinen Schule ggf. nicht

mehr geleistet werden könne und somit die Sonderpädagogik beziehungsweise Förderschule eingreifen müsse (Kottmann, 2006; Powell, 2004). Hier zeigt sich auch der zirkuläre Charakter der KMK-Definition: Ein SPF liegt vor, wenn eine Schülerin oder ein Schüler ohne sonderpädagogische Unterstützung nicht am Unterricht der allgemeinen Schule teilnehmen kann. Durch diesen Zirkelschluss verstärke sich die Zuständigkeit der Sonderpädagogik für die Schülergruppe mit SPF (Kottmann, 2006).

Zweitens wird die Unterscheidung zwischen sonderpädagogischem Förderbedarf und allgemeinem Förderbedarf als problematisch erachtet (Eberwein, 1996). Jede Schülerin beziehungsweise jeder Schüler habe einen Förderbedarf, der individuell verschieden ist (Kottmann, 2006). Die fließende Grenze zwischen sonderpädagogischem und regulärem Förderbedarf erschwere die Feststellungsdiagnostik eines SPF (Eberwein, 1996).

Drittens wird am Begriff des SPF kritisiert, dass er sich nicht an gängigen Klassifikationssystemen orientiere (Powell & Pfahl, 2012), wie zum Beispiel dem DSM-V (Diagnostic and Statistical Manual of Mental Disorders, 5. Auflage; American Psychiatric Association, 2013) oder der ICD-10 (International Statistical Classification of Diseases and Related Health Problems, 10. Revision; World Health Organization, WHO, 2016). Dies könne die Schaffung bundesweit einheitlicher Diagnosekriterien zur Feststellung eines SPF erschweren (Bos et al., 2010; Grosche, 2015).

Viertens wird kritisiert, dass die Feststellung eines SPF mit einer Kategorisierung verbunden ist, die für Kinder und Jugendliche stigmatisierend wirken könne (Biewer, 2010; Grosche, 2015). Deshalb fordern einige Forscherinnen und Forscher, auf eine Kategorisierung von Schülerinnen und Schülern mit den Förderschwerpunkten *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* zu verzichten (Hinz, 2010). So könnte die Feststellung eines SPF in diesen Bereichen mit geringeren Leistungserwartungen einhergehen (Gomolla & Radkte, 2009; Powell & Pfahl, 2012), welche von den Schülerinnen und Schülern internalisiert werden und zu einem geringeren Selbstwertgefühl beitragen könnten (Knigge, 2017; Pfahl, 2011). Wocken (2012) verweist darauf, dass eine individualisierte Förderung gerade nicht auf Kategorien basiere, da durch Kategorien individuelle Unterschiede innerhalb der Schülergruppe mit einem bestimmten Förderschwerpunkt nicht ausreichend berücksichtigt werden können.

2.3.2 Verfahren der Feststellung eines sonderpädagogischen Förderbedarfs

2.3.2.1 Verfahren zur Feststellung eines SPF

Verfahren zur Feststellung eines SPF werden durch bundesweite Richtlinien zu Definitionen und Förderorten für Schülerinnen und Schüler mit SPF geregelt, die die Bundesländer in ihren Schulgesetzen interpretieren (s. KMK, 1994, 2011; Sälzer, Gebhardt, Müller & Pauly, 2015). Dabei unterscheiden sich die Bundesländer in verschiedenen Aspekten des Feststellungsverfahrens (z. B. Zeitpunkt, Diagnosekriterien, Elternmitspracherecht bei der Wahl des Förderortes). Jedoch bestehen auch Gemeinsamkeiten. So kommen bei der Feststellung eines SPF in allen Bundesländern wissenschaftlich erprobte und validierte Diagnoseverfahren zum Einsatz (Sälzer et al., 2015). Am Feststellungsverfahren sind die jeweilige Schule, die Erziehungsberechtigten und die Schulaufsichtsbehörde (bzw. vergleichbare Einrichtungen) beteiligt. In den meisten Fällen beantragt die Schule beziehungsweise Schulaufsichtsbehörde die Einleitung eines Feststellungsverfahrens (Sälzer et al., 2015). In vielen Bundesländern können auch die Erziehungsberechtigten oder Schülerinnen beziehungsweise Schüler ab dem 14. beziehungsweise 18. Lebensjahr die Einleitung eines Feststellungsverfahrens beantragen (Sälzer et al., 2015).

Das Feststellungsverfahren kann vor der Einschulung, während der Schuleingangsuntersuchung, oder auch nach der Einschulung eingeleitet werden (Sälzer et al., 2015), wobei zunächst Hinweise aus dem Schulalltag herangezogen werden (Petermann & Petermann, 2006). Dazu gehören unter anderem deutliche schulische Minderleistungen (z. B. zweimaliges Sitzenbleiben), Verzögerungen in der expressiven und rezeptiven Sprachentwicklung und Verhaltensauffälligkeiten (Kottmann, 2006; Petermann & Petermann, 2006; Sälzer et al., 2015). Die Feststellung eines SPF erfolgt in einigen Bundesländern auch im Anschluss an die Schuleignungsuntersuchung (Sälzer et al., 2015), die in den meisten Bundesländern verpflichtend ist und unter anderem der Feststellung eines Förderbedarfs bei Kindern im Vorschulalter dient.

Die Bundesländer unterscheiden sich darin, ob eine frühe oder späte Feststellung eines SPF in den Bereichen *Lernen*, *Sprache* oder *emotionale und soziale Entwicklung* angestrebt wird (Gresch et al., 2014). Manche Bundesländer (z. B. Berlin, Bremen, Schleswig-Holstein) verzichten auf Feststellungsverfahren in den ersten Schuljahren (Sälzer et al., 2015). Dabei wird argumentiert, dass die Diagnose eines SPF mit Unsicherheiten verbunden ist, da sich

Lernschwierigkeiten erst im Laufe der Schulkarriere manifestieren (Gresch et al., 2014). Eine frühe Feststellung könne zudem stigmatisierend wirken und sich negativ auf die Leistungsentwicklung der diagnostizierten Schülerinnen und Schüler auswirken (Dunn, 1968). In diesen Bundesländern werden alle Schülerinnen und Schüler in den ersten Grundschuljahren gemeinsam unterrichtet und sonderpädagogische Ressourcen lerngruppenbasiert zugewiesen. Ein Feststellungsverfahren wird nach den ersten Schuljahren eingeleitet, wenn eine Schülerin beziehungsweise ein Schüler mit den in allgemeinen Schule zur Verfügung stehenden sonderpädagogischen Ressourcen nicht ausreichend gefördert werden kann (Sälzer et al., 2015). Eine späte Feststellung wird auch in Schulen präferiert, die inklusiv ausgerichtet sind und in denen sonderpädagogische Ressourcen zur Verfügung stehen (z. B. Rügener Inklusionsmodell in Mecklenburg-Vorpommern). In diesen Schulen wird zunächst mit präventiven Maßnahmen gearbeitet (Gresch et al., 2014). In anderen Bundesländern (z. B. Sachsen) wird die Feststellung eines SPF am Anfang der Schullaufbahn eingeleitet, um frühzeitig zusätzliche Unterstützungsangebote und Ressourcen bereitstellen zu können (Gresch et al., 2014). Die Variabilität in den Feststellungsverfahren bedeutet für die vorliegende Dissertation, dass bei der Schätzung von Schularteffekten Unterschiede zwischen den Bundesländern berücksichtigt werden müssen.

2.3.2.2 Diagnostische Instrumente zur Feststellung eines SPF

Das Feststellungsverfahren eines SPF lässt sich den Bereichen der Selektionsdiagnostik und der Platzierungsdiagnostik zuordnen. Ziel dieser Diagnostik ist es, für Schülerinnen und Schüler einen passenden Bildungsweg (Selektion) beziehungsweise einen passenden Lernort (Platzierung) auszuwählen (Marx & Lenhard, 2011). Im Feststellungsverfahren kommen unterschiedliche Methoden zum Einsatz (Kany & Schöler, 2009). Dazu gehören Elternbefragungen, Expertenurteile, Verhaltensbeobachtungen und psychometrische Tests, die zur Beurteilung der kognitiven, motivationalen und sozio-emotionalen Ausgangslagen von Kindern herangezogen werden (Petermann & Petermann, 2006).⁴

⁴ Außerdem schreibt das Neunte Buch Sozialgesetzbuch ab dem 1. Januar 2018 vor, dass sich die Ermittlung eines individuellen Bedarfs für Leistungen der Eingliederungshilfe, zu denen auch Leistungen zur Teilhabe an Bildung gehören, an der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) orientieren muss (§ 118 Abs. 1 SGB IX i. d. F. d. Bundesteilhabegesetzes).

Elternbefragungen und Expertenurteile dienen einer differenzierten Kind-Umwelt-Analyse (Kany & Schöler, 2009). Dabei sollen Risiko- und Schutzfaktoren in der Umwelt des Kindes ermittelt werden, um Ressourcenzuweisungen mit bereits vorhandenen Ressourcen, aber auch mit Risikofaktoren abstimmen zu können (Kottmann, 2006; Sälzer et al., 2015). Darüber hinaus werden bei der Feststellung eines SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* Expertinnen und Experten einbezogen, um organische Beeinträchtigungen, unter anderem im Bereich der Sinneswahrnehmungen, aus fachärztlicher Perspektive ausschließen zu können. Verhaltensbeobachtungen werden überwiegend von Lehrkräften durchgeführt und dienen dazu, das Lern- und Leistungsverhalten und das Sozialverhalten der Kinder und Jugendlichen einzuschätzen (Petermann & Petermann, 2006). Diese Beobachtungsverfahren sind selten standardisiert und haben eher explorativen Charakter. Allerdings existieren auch systematische Beobachtungsverfahren, die in Förderprogramme eingebunden sind (z. B. Sprachförderung nach Frank, Kirschhock & Martschinke, 2003). Bei der Einschätzung des Lern- und Sozialverhaltens kommen auch Beurteilungsskalen zum Einsatz, wie zum Beispiel die *Diagnostischen Einschätzskalen zur Beurteilung des Entwicklungsstandes und der Schulfähigkeit* (DES; Barth, 2016) und der *Verhaltensbeurteilungsbogen Schule* (VBS-L; Petermann, Menzel & Eckert, 2003). Herausforderungen bei der Feststellung eines SPF bestehen darin, dass der Einsatz informeller Verfahren (z. B. Beurteilungen, unstandardisierte Beobachtungen) zu einer geringeren Objektivität führen könne, wodurch das gesamte Verfahren anfällig für Beobachtereffekte ist (z. B. *Halo-Effekte*, *Primacy-Effekt*, *Confirmation Bias*; Kany & Schöler, 2009; Kottmann, 2006).

Außerdem kommen bei der Feststellung eines SPF auch psychologische Tests zum Einsatz. Dazu zählen unter anderem Tests zur Gedächtnisentwicklung (z. B. *Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12*; Hasselhorn et al., 2012), zur visuellen und auditiven Wahrnehmung (z. B. *Frostigs Entwicklungstest der visuellen Wahrnehmung*; Büttner, Dacheneder, Schneider & Weyer, 2008), zur kognitiven Leistungsfähigkeit (z. B. *Kaufman-Assessment-Battery for Children*; Kaufman & Kaufman, 2009) und zu sprachlichen Fähigkeiten (z. B. *Heidelberger Sprachentwicklungstest*; Grimm & Schöler, 1991). Die Auswahl der spezifischen diagnostischen Verfahren ist dabei von der genauen Fragestellung geleitet, insbesondere welcher Förderschwerpunkt vermutet wird (Kany & Schöler, 2009).

Intelligenztests beziehungsweise Tests zu kognitiven Fähigkeiten zählen zu den am häufigsten eingesetzten diagnostischen Instrumenten bei der Feststellung eines SPF im Bereich *Lernen*. Sie dienen der Differenzierung zwischen den Förderschwerpunkten *Lernen* und *geistige*

Entwicklung (KMK, 1999). In beiden Förderschwerpunkten sind deutlich unterdurchschnittliche Testwerte ein zentrales Kriterium für die Diagnose eines SPF (Kretschmann, 2006). Eine deutlich unterdurchschnittliche Intelligenztestleistung bezieht sich im Förderschwerpunkt *Lernen* auf mindestens 1–1.5 Standardabweichungen von der jeweiligen Altersgruppe (z. B. Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, 2012), während die Testleistung für die Feststellung eines SPF im Bereich *geistige Entwicklung* in Anlehnung an die Klassifikationssysteme DSM-V (American Psychiatric Association, 2013) und ICD-10 (WHO, 2016) mindestens 2 Standardabweichungen betragen sollte. Kritisch angemerkt wird jedoch, dass zwischen den Bundesländern unterschiedliche Grenzwerte existieren, was die Vergleichbarkeit der Schülergruppen mit diesen Förderschwerpunkten einschränkt (Kretschmann, 2006). Bei der Diagnose eines SPF in den Bereichen *Sprache* und *emotionale und soziale Entwicklung* werden Intelligenztests hingegen seltener eingesetzt. Sie dienen in diesen Bereichen lediglich der Differenzialdiagnose.

Der Einsatz von Intelligenztests wird aus pädagogischer Sicht dahingehend kritisiert, dass sich aus den Testergebnissen keine direkten Förderansätze ableiten lassen (Kottmann, 2006; Werning & Lichtblau, 2012). Die Ableitung von konkreten Interventionen ist jedoch nicht Aufgabe der Eingangsdiagnose und sollte sich vielmehr dem Feststellungsverfahren anschließen (s. a. Petermann & Petermann, 2006). Zum anderen wird kritisiert, dass durch die prominente Stellung von Intelligenztests bei der Feststellung eines SPF (insbesondere im Bereich *Lernen*) kognitive Aspekte der Beeinträchtigung zu stark in den Vordergrund gestellt werden (Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, 2012). Es wird dennoch empfohlen, mit einem „Breitbandtest“ (z. B. Intelligenztest) zu beginnen und anschließend spezifischere Tests zur Diagnose des Förderschwerpunktes einzusetzen (Kany & Schöler, 2009).

Neben der Eingangsdiagnose zur Feststellung des Förderschwerpunktes und des Förderortes können drei weitere Bereiche der sonderpädagogischen Diagnostik unterschieden werden, die sich dem Feststellungsverfahren anschließen (Petermann & Petermann, 2006). Dazu zählt erstens die Förderdiagnose, bei der ein Förderplan unter Berücksichtigung individueller Entwicklungsprofile erstellt wird. Dabei werden Förderziele, Fördermethoden und die zeitlichen Rahmenbedingungen festgelegt (Petermann & Petermann, 2006). Ziel der Förderdiagnose ist auch, Impulse für die Unterrichtsgestaltung abzuleiten, die Schülerinnen und Schüler mit SPF bei der Erreichung der Förderziele unterstützen können. Zweitens werden diagnostische Verfahren zur systematischen Dokumentation der Lernentwicklung von

Kindern und Jugendlichen mit SPF eingesetzt (Autorengruppe Bildungsberichterstattung, 2014). Dabei wird evaluiert, welche Lernfortschritte und -schwierigkeiten im Prozess der Förderung auftreten und welche Rolle die Lernumgebung spielt (Petermann & Petermann, 2006). Drittens erfolgt im Rahmen der Verlaufskontrolle eine Fortschreibungsdiagnose (Petermann & Petermann, 2006). Dabei wird nach Abschluss der sonderpädagogischen Förderung ein Resümee gezogen und gegebenenfalls eine Anpassung des Förderplans vorgenommen. Ein SPF wird solange fortgeschrieben, bis das Förderziel erreicht ist (Petermann & Petermann, 2006). Im Gegensatz zu den diagnostischen Verfahren bei der Feststellung eines SPF zeichnet sich die sonderpädagogische Förderdiagnostik durch eine stärkere kriteriale Orientierung aus. Sie dient also nicht der vergleichenden Beschreibung und Differenzierung von Schülerinnen und Schülern, sondern der Unterstützung individueller Lern- beziehungsweise Förderziele (Petermann & Petermann, 2006).

2.3.2.3 Zuweisung zu einer Beschulungsart

Nach der Feststellung eines SPF schließt das Verfahren mit einer Empfehlung zum Förderort. Diese Entscheidung obliegt der jeweiligen Schulaufsicht (Sälzer et al., 2015). Die Schulaufsichtsbehörde bezieht in die Entscheidungsfindung die Schule der Schülerinnen und Schüler sowie deren Lehrkräfte und Erziehungsberechtigte ein (KMK, 1999). Bei der Wahl des Förderortes werden zunächst der Umfang des SPF beziehungsweise der Schweregrad und die Komplexität der Beeinträchtigung berücksichtigt (Petermann & Petermann, 2006). Liegt nur ein geringer Förderbedarf vor, werden Schülerinnen und Schüler eher in allgemeinen Schulen unterrichtet (Gresch et al., 2014; Sälzer et al., 2015). Darüber hinaus beeinflussen die Stellungnahme der Eltern, die Verfügbarkeit sonderpädagogischen Personals und spezieller Lehr- und Lernmittel sowie die Stellungnahme der Maßnahmenträger die Entscheidung über den Förderort (KMK, 1999). Zusätzlich beeinflussen strukturelle Merkmale, wie die Verfügbarkeit sonderpädagogischer Fördereinrichtungen am Wohnort, den Entscheidungsprozess.

Neben förderschwerpunktübergreifenden Richtlinien hat die KMK förderschwerpunktspezifische Empfehlungen zum Förderort erstellt (KMK, 1998, 1999, 2000). Diese halten fest, dass bei Schülerinnen und Schülern, die einen SPF im Bereich *Lernen* diagnostiziert bekommen, zunächst eine Beschulung in allgemeinen Schulen anzustreben ist (KMK, 1999). Nur wenn im Einzelfall keine angemessene Förderung (bezogen auf personelle, räumliche und sächliche Voraussetzungen) ermöglicht werden kann, sollte eine Beschulung in der Förderschule

gewählt werden. Bei Schülerinnen und Schülern mit dem Förderschwerpunkt *Sprache* wird der Fokus auf vorbeugende Maßnahmen und Frühförderung gelegt. Dabei ist die Förderschule in diesem Förderschwerpunkt als Durchgangsschule konzipiert (KMK, 1998, 2015). Sobald die sprachlichen Kompetenzen der betroffenen Schülerinnen und Schüler eine erfolgreiche Teilnahme am Regelunterricht ermöglichen, sollte eine (Rück-)Überweisung auf eine allgemeine Schule erfolgen. Im Gegensatz zum Förderschwerpunkt *Lernen* ist hier eine frühe Aufnahme an der Förderschule intendiert, da sprachliche Defizite als elementare Hindernisse für den weiteren Lernverlauf angesehen werden, denen zum frühestmöglichen Zeitpunkt entgegengewirkt werden sollte (KMK, 1998). Auch für Kinder und Jugendliche mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* ist die Förderschule als Durchgangsschule mit dem Ziel einer schnellen Rückführung in die allgemeine Schule konzipiert (KMK, 2015). Allerdings wurden bisher keine objektiven Kriterien für die Entscheidung über den Förderort entwickelt, was dazu führen könnte, dass Förderschulüberweisungen stark durch die an der Feststellung beteiligten Akteure beeinflusst werden (Mand, 2003).

Unterschiede zwischen den Bundesländern bestehen auch darin, wie stark die Erziehungsberechtigten bei der Entscheidung über den Förderort einbezogen werden (Sälzer et al., 2015). Nach der Feststellung eines SPF haben die Eltern bundesweit ein Mitspracherecht bei der Entscheidung über den Förderort, aber die Entscheidung trifft in den meisten Ländern die Schulaufsichtsbehörde (Sälzer et al., 2015). In manchen Bundesländern ist der Elternwunsch nicht bindend (z. B. Brandenburg); in anderen Bundesländern (z. B. Berlin, Saarland) dagegen kann die Schulaufsichtsbehörde kaum gegen den Willen der Eltern beziehungsweise die Empfehlung des Förderausschusses entscheiden (Sälzer et al., 2015). Außerdem unterscheiden sich die Bundesländer darin, zu welchem Zeitpunkt Schülerinnen und Schüler mit SPF an eine Förderschule überwiesen werden können (Gresch et al., 2014). In Bundesländern (z. B. Berlin), in denen auf eine Feststellung eines SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* in den ersten Schuljahren verzichtet wird, ist die Überweisung auf eine Förderschule zunächst prinzipiell nicht vorgesehen (Gresch et al., 2014). In manchen Bundesländern beginnen Förderschulen mit dem Förderschwerpunkt *Lernen* (z. B. Berlin, Mecklenburg-Vorpommern, Schleswig-Holstein) erst mit der dritten Jahrgangsstufe (Gresch et al., 2014).

Abschließend muss festgehalten werden, dass die vorliegende Darstellung zur Feststellung eines SPF und zur Entscheidung über den Förderort auf länderübergreifenden Empfehlungen der KMK beziehungsweise auf den Schulgesetzen der Bundesländer basiert. Die Ausgestal-

tung dieser Empfehlungen beziehungsweise Schulgesetze variiert, so dass Diskrepanzen zwischen den festgelegten Richtlinien und der tatsächlichen Implementation auftreten können (Kottmann, 2006; Wrase, 2015). So wird beispielsweise von Seiten der KMK ausdrücklich darauf hingewiesen, dass die Feststellung eines SPF und die Entscheidung über den Förderschwerpunkt in angemessenen zeitlichen Abständen überprüft werden sollten. In der Praxis sind regelmäßige Überprüfungen oder Rücküberweisungen von der Förderschule in die allgemeine Schule jedoch die Ausnahme und das Konzept der Förderschule als Durchgangsschule eher ein konzeptueller Vorschlag als bildungspolitische Realität (Autorengruppe Bildungsberichtserstattung, 2012; Dietze, 2012; Kottmann, 2006). Im Einzelfall könnten zudem fehlende sonderpädagogische Ressourcen in allgemeinen Schulen zur Empfehlung für eine Beschulung in einer Förderschule führen (Sälzer et al., 2015).

2.3.3 Situation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Deutschland

2.3.3.1 Beschreibung der Schülergruppe mit SPF auf Grundlage amtlicher Schulstatistiken

Die KMK gibt alle zwei Jahre eine Dokumentation zur Situation von Schülerinnen und Schülern mit SPF in allgemeinbildenden Schulen von der ersten bis zehnten Jahrgangsstufe heraus (KMK, 2016). Dabei werden alle Kinder und Jugendlichen gezählt, die sonderpädagogisch in Förderschulen oder allgemeinen Schulen gefördert werden. Die Darstellung basiert auf amtlichen Schulstatistiken der Kultusministerien der Bundesländer. Die Situation von Kindern und Jugendlichen mit SPF wird in den Schulstatistiken anhand von drei Quoten beschrieben: Förderquote, Integrations- beziehungsweise Inklusionsquote und Förderschulbesuchsquote (Malecki, 2014). Die *Förderquote* beschreibt den prozentualen Anteil von Schülerinnen und Schülern mit SPF an der Gesamtschülerzahl von der ersten bis zehnten Jahrgangsstufe. Sie entspricht der Summe aus Förderschulbesuchsquote und Integrations- beziehungsweise Inklusionsquote. Die *Inklusionsquote* beschreibt den prozentualen Anteil an Schülerinnen und Schülern mit SPF, die eine allgemeine Schule anstatt einer Förderschule besuchen. Die *Förderschulbesuchsquote* gibt den Anteil der Kinder und Jugendlichen in Förderschulen an der gesamten Schülerpopulation an. Sie wird auch im internationalen Vergleich verwendet, um das Ausmaß an Separation beziehungsweise Segregation im Bildungssystem einzuordnen (Biermann & Powell, 2014).

Im Schuljahr 2013/2014 wurde bei circa 7 Prozent (ca. 510 000 Schülerinnen und Schüler) der gesamten Schülerpopulation der ersten bis zehnten Jahrgangsstufe ein sonderpädagogischer Förderbedarf festgestellt. Im Zeitverlauf zeigte sich zwischen 1999 und 2014 ein kontinuierlicher Anstieg des Anteils an Schülerinnen und Schülern mit SPF an der gesamten Schülerpopulation (Förderquote). Die absolute Größe der Schülergruppe blieb in diesem Zeitraum allerdings relativ konstant, bei gleichzeitig sinkender Gesamtschülerzahl. Somit ist der Anstieg der Förderquote vor allem auf den Rückgang der Schülerzahl insgesamt zurückzuführen.

Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* bilden die größte Schülergruppe mit SPF. Allerdings nahm der relative Anteil dieser Gruppe in den letzten 15 Jahren kontinuierlich ab (Abbildung 1): Während im Schuljahr 1998/99 noch circa 55 Prozent der Schülergruppe mit SPF den Förderschwerpunkt *Lernen* erhielten, lag dieser Anteil im Schuljahr 2013/2014 bei circa 38 Prozent. Die zweitgrößten Schülergruppen mit SPF waren im Schuljahr 2014/15 Schülerinnen und Schüler mit dem Förderschwerpunkt *geistige Entwicklung* und Schülerinnen und Schüler mit dem Förderschwerpunkt *emotionale und soziale Entwicklung*. Der Anteil der Schülergruppe mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* hat sich in den letzten 15 Jahren mehr als verdoppelt (von 7% im Schuljahr 1998/99 auf 16% im Schuljahr 2013/2014). Der Förderschwerpunkt *Sprache* wurde im Schuljahr 2013/2014 bei knapp 11 Prozent der Schülerinnen und Schülern mit SPF diagnostiziert. Dieser Anteil ist in den letzten 15 Jahren relativ stabil geblieben (Abbildung 1).

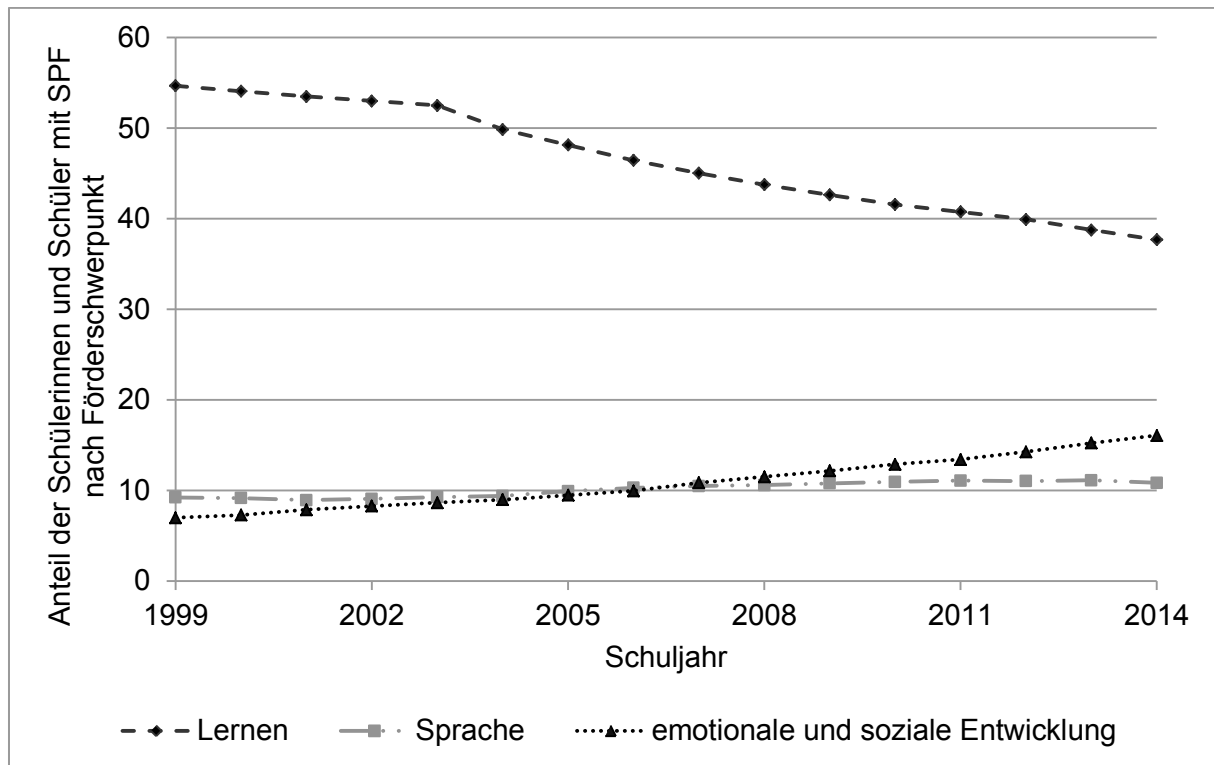


Abbildung 1. Prozentualer Anteil der Schülerinnen und Schüler mit den Förderschwerpunkten Lernen, Sprache und emotionale und soziale Entwicklung an allen Schülerinnen und Schülern mit SPF (KMK, 2016).

In den letzten 15 Jahren stieg der Anteil an Schülerinnen und Schülern mit SPF in allgemeinen Schulen kontinuierlich an (Inklusionsquote). Wurden im Schuljahr 1998/1999 nur circa 12 Prozent der Schülerinnen und Schüler mit SPF in allgemeinen Schulen unterrichtet, lag dieser Anteil im Schuljahr 2013/2014 schon bei circa 34 Prozent (KMK, 2016). Dieser Trend lässt sich auch für die drei Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* feststellen (Abbildung 2). Seit der Verabschiedung der UN-BRK (2008) hat sich der Trend hin zu einer inklusiven Beschulung also verstärkt. Trotzdem wird die Mehrheit der Kinder und Jugendlichen mit SPF weiterhin in Förderschulen unterrichtet. Einzig Schülerinnen und Schüler mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* werden mehrheitlich in allgemeinen Schulen unterrichtet. Die Inklusionsquote ist zudem mit zunehmender Jahrgangsstufe niedriger (Weishaupt, 2016): Im Primarbereich ist der Anteil an Kindern mit SPF in allgemeinen Schulen deutlich höher als in der Sekundarstufe (Autorengruppe Bildungsberichterstattung, 2016). Ohne den Förderschwerpunkt *geistige Entwicklung* beträgt der Anteil integrativ beziehungsweise inklusiv unterrichteter Schülerinnen und Schüler mit SPF in der Grundschule rund 44 Prozent, während der Anteil in der Sekundarstu-

fe I bei insgesamt 23 Prozent liegt (Autorengruppe Bildungsberichterstattung, 2014; Sälzer et al., 2015).

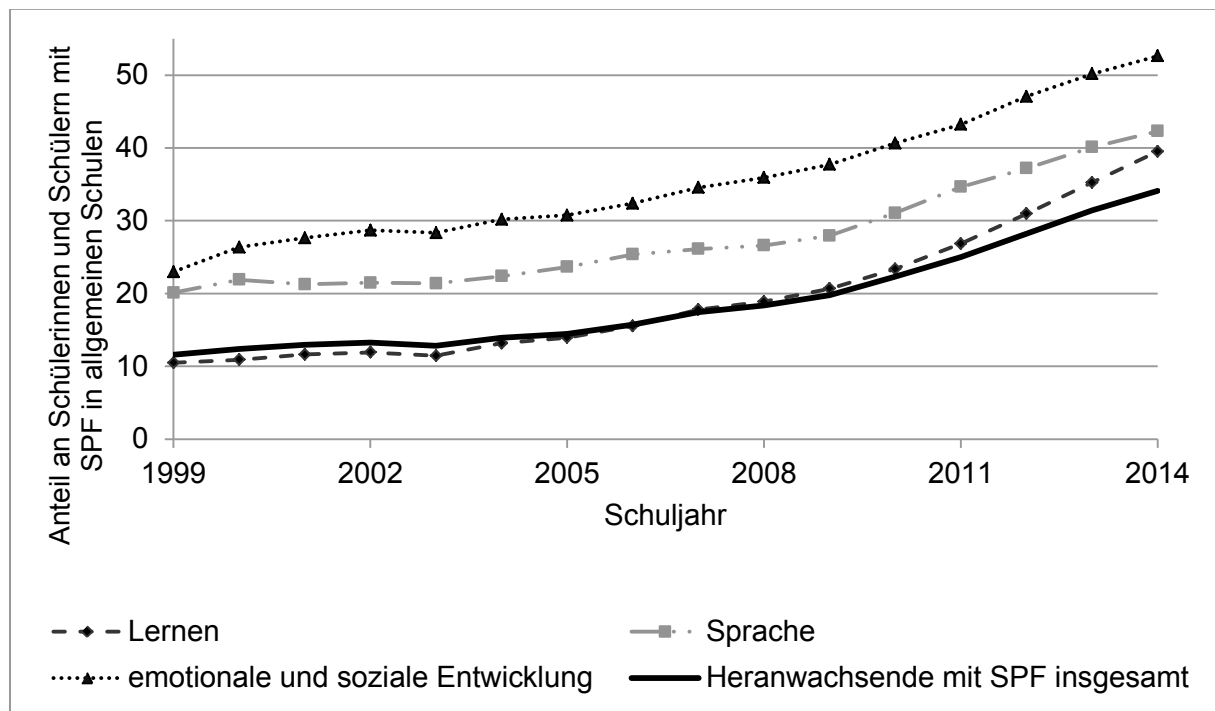


Abbildung 2. Entwicklung der Inklusionsquote nach Förderschwerpunkt vom Schuljahr 1999/2000 bis zum Schuljahr 2014/2015 (KMK, 2016).

Auch wenn immer mehr Schülerinnen und Schüler mit SPF in allgemeinen Schulen unterrichtet werden, ist der Anteil an Schülerinnen und Schülern, die in Förderschulen unterrichtet werden, gemessen an der gesamten Schülerpopulation in den letzten 15 Jahren relativ konstant geblieben (KMK, 2016). Diese konstante Förderschulbesuchsquote wird dahingehend interpretiert, dass die Verabschiedung der UN-BRK (2008) noch nicht zu tiefgreifenden Veränderungen im deutschen Bildungssystem geführt hat (Werning, 2014; Wrase, 2016). Vielmehr wird das hoch differenzierte Förderschulsystem mit speziell ausgebildeten Lehrkräften noch aufrechterhalten (Gomolla & Radtke, 2009; Kottmann, 2006). Insgesamt sind circa 10 Prozent aller allgemeinbildenden Schulen Förderschulen, wobei dieser Anteil zwischen den Bundesländern stark variiert (von 5.4% in Bremen bis 13.6% in Mecklenburg-Vorpommern; Autorengruppe Bildungsberichterstattung, 2016). Die Anzahl der Förderschulen ist zwischen 2006 und 2014 weitgehend konstant geblieben, auch wenn in den ostdeutschen Ländern ein leichter Rückgang beobachtbar ist (Autorengruppe Bildungsberichterstat-

tung, 2016). Trotz inklusiver Bemühungen existiert also weiterhin ein großes Angebot an Förderschulen (Autorengruppe Bildungsberichterstattung, 2016). Dass dieses Angebot weiterhin genutzt wird, zeigt sich auch darin, dass der Anteil an Schülerinnen und Schülern, die direkt in die Förderschule eingeschult wurden, seit 2003 konstant bei circa 3 Prozent liegt (Autorengruppe Bildungsberichterstattung, 2016). Allerdings verzichten einige Bundesländer (insbesondere die Stadtstaaten) nun auf die Aufrechterhaltung separater Förderschulen für die Förderschwerpunkte *Lernen* und *emotionale und soziale Entwicklung*, was sich auch in einer Verringerung der Einschulungsquoten in Förderschulen mit diesen Förderschwerpunkten widerspiegelt (Autorengruppe Bildungsberichterstattung, 2016).

Auf Länderebene bestehen deutliche Unterschiede in der sonderpädagogischen Förderung (Abbildung 3; Autorengruppe Bildungsberichterstattung, 2016). Zum einen variiert der Anteil an Schülerinnen und Schülern mit SPF an der gesamten Schülerpopulation (Förderquote). Während deutschlandweit bei circa 7 Prozent aller Schülerinnen und Schüler ein SPF diagnostiziert wurde, schwankt dieser Anteil zwischen 5.6 Prozent in Rheinland-Pfalz und 10.6 Prozent in Mecklenburg-Vorpommern. Deutliche Unterschiede zwischen den Bundesländern zeigen sich insbesondere im Anteil der Kinder und Jugendlichen, bei denen ein SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* festgestellt wird (eigene Berechnungen auf Grundlage von KMK, 2016). Diese Befunde können darauf zurückgeführt werden, dass die Diagnosekriterien für die Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* weniger konsistent sind (Kany & Schöler, 2009; Sälzer et al., 2015). Darüber hinaus unterscheiden sich die Bundesländer auch im Anteil der Schülerinnen und Schüler mit SPF, die in allgemeinen Schulen unterrichtet werden. Die Inklusionsquote variiert zwischen 22.8 Prozent in Hessen und 76.9 Prozent in Bremen (Abbildung 3).

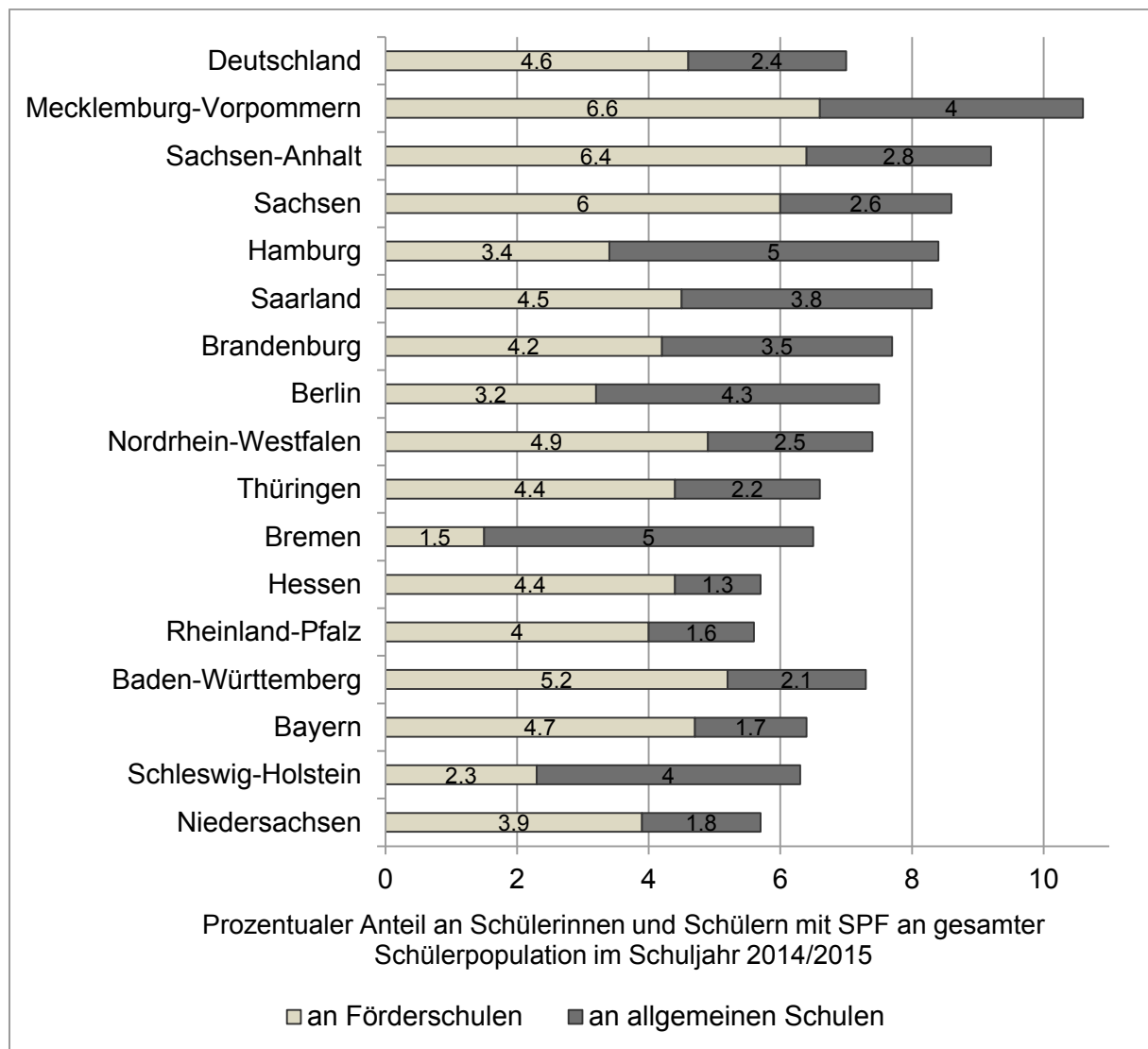


Abbildung 3. Sonderpädagogische Förderung im Schuljahr 2014/2015 in den Bundesländern (Autorengruppe Bildungsberichterstattung, 2016; S. 81).

Zusammenfassend lässt sich zum Stand der sonderpädagogischen Förderung festhalten, dass trotz der Veränderungen seit der Verabschiedung der UN-BRK (2008) in Deutschland weiterhin ein separates Förderschulsystem besteht, das für die Mehrheit der Schülerinnen und Schüler mit SPF zuständig ist. Die Bildungseinrichtungen scheinen sich momentan in einer Umbruchphase zu befinden, wobei die Umgestaltung der Lernumwelten in Einklang mit der UN-BRK (2008) im Primarbereich schon weiter fortgeschritten zu sein scheint als in der Sekundarstufe (Autorengruppe Bildungsberichterstattung, 2014; Baumert, 2016a). Diese Umgestaltung bedeutet einen Traditionsbruch mit dem differenzierten sonderpädagogischen Fördersystem und erfordert sowohl organisatorische und strukturelle Veränderungen als auch eine kulturelle Neuorientierung (Baumert, 2016a). In einigen Bundesländern (z. B. Bremen,

Schleswig-Holstein), in denen schon früh inklusive Ansätze erprobt wurden, ist diese Umgestaltung schon weiter fortgeschritten (Biermann & Powell, 2014; Preuss-Lausitz, 2001). Insbesondere in Bundesländern mit einem ausgeprägten Förderschulsystem (z. B. Bayern) werden Förderschulen trotz der sinkenden Schülerzahl insgesamt weiterhin aufrechterhalten (s. a. Kapitel 2.3.3.3).

Ein Vorteil amtlicher Schulstatistiken liegt in der Vollständigkeit der Daten, sie geben also ein repräsentatives Bild der Situation von Kindern und Jugendlichen mit SPF in Deutschland wieder. Für die vorliegende Dissertation liefern amtliche Schulstatistiken Hinweise darauf, dass Unterschiede zwischen den Bundesländern im Anteil der Schülerinnen und Schüler mit SPF in allgemeinen Schulen bei der Schätzung von Schularteffekten berücksichtigt werden sollten.

Die zentrale Herausforderung bei der bundesweiten Auswertung von Schulstatistiken resultiert daraus, dass diese in der Verantwortlichkeit der Bundesländer liegen und von den dortigen Kultusbehörden erhoben werden (Malecki, 2014). Zwischen den Bundesländern gibt es keine einheitlichen Kriterien dafür, welche Kinder und Jugendlichen zur Schülergruppe mit SPF gezählt werden (Malecki, 2014). Zwar wird in den Veröffentlichungen der KMK (2016) beschrieben, dass die erhaltene sonderpädagogische Förderung ausschlaggebend für eine Berücksichtigung in den Schulstatistiken ist, dieses Kriterium wird jedoch in den Bundesländern nicht einheitlich operationalisiert. So wird in manchen Bundesländern ausschließlich die erhaltene Förderung als Kriterium genutzt, in anderen Bundesländern hingegen die amtliche Feststellung eines SPF (Gresch et al., 2014). Ferner kommt es auch vor, dass beide Kriterien zur statistischen Erfassung von Kindern und Jugendlichen mit SPF verwendet werden (Gresch, et al., 2014). Dadurch ist die Vergleichbarkeit der Statistiken zwischen den Ländern nicht immer gewährleistet.

Außerdem besteht eine wichtige Limitation amtlicher Schulstatistiken darin, dass keine detaillierten Informationen zur Lernumgebung und Ausgestaltung des gemeinsamen Unterrichts in allgemeinen Schulen erfasst werden. Je nach Bundesland stehen unterschiedliche Konzepte hinter den statistisch erfassten Maßnahmen der schulischen Inklusion (Malecki, 2014). Unterschiede bestehen zum Beispiel in der Differenzierung zwischen zielgleichem und ziendifferentem Lernen, im Umfang beziehungsweise in der Dauer der sonderpädagogischen Förderung und in der Organisation von Schulklassen mit Schülerinnen und Schülern mit SPF in allgemeinen Schulen (Malecki, 2014). Eine Kooperationsklasse in einer allgemeinen Schule, in der ausschließlich Schülerinnen und Schüler mit SPF unterrichtet werden, könnte

zum Beispiel in der amtlichen Schulstatistik als inklusives Setting gezählt werden, obwohl in solchen Klassen Schülerinnen und Schüler mit SPF nicht gemeinsam mit Schülerinnen und Schülern ohne SPF lernen (Wrase, 2015).

2.3.3.2 Das Ressourcen-Etikettierungs-Dilemma

Die Auswertung amtlicher Schulstatistiken zeigt, dass die Förderquote in den letzten 15 Jahren kontinuierlich gestiegen ist (KMK, 2016). Für diesen steigenden Anteil an Schülerinnen und Schülern mit SPF im Vergleich zur gesamten Schülerpopulation werden unterschiedliche Gründe diskutiert. Eine mögliche Erklärung wird als *Ressourcen-Etikettierungs-Dilemma* bezeichnet (Bleidick, 1988; Füssel & Kretschmann, 1993; Wocken, 1996). Das Dilemma besteht darin, dass die Zuweisung von Ressourcen für Schülerinnen und Schüler mit SPF an die Feststellung eines SPF geknüpft ist (Wocken, 1996). Somit ist der Erhalt sonderpädagogischer Unterstützung mit der etikettierenden Feststellung eines SPF verbunden. Schulen erhielten eine positive Verstärkung in Form von höherer Mittelzuweisung dafür, eine große Anzahl von Schülerinnen und Schülern mit SPF zu diagnostizieren (Füssel & Kretschmann, 1993). Insbesondere allgemeine Schulen könnten verstärkt Feststellungsverfahren einleiten, um Personal beziehungsweise finanzielle Mittel zu erhalten. Die Feststellung eines SPF dient dabei sowohl der Sicherung von Ressourcen als auch der Legitimation der Forderung nach sonderpädagogischer Unterstützung. Diese Verschränkung führt nach Wocken (1996) zum Anstieg der Förderquote. Eine weitere Erklärung ist, dass die Folgen einer SPF-Diagnose durch inklusive Bemühungen als nicht mehr so gravierend und deterministisch gesehen werden und die Schwellen für eine Feststellung daher sinken (Powell & Pfahl, 2012).

Dass sonderpädagogische Ressourcen erst nach der formalen Feststellung eines SPF bereitgestellt werden, könnte negative Folgen für die betroffenen Schülerinnen und Schüler und Lehrkräfte haben (Wocken, 1996). Für Kinder und Jugendliche mit vorübergehenden schulischen Problemen oder Teilleistungsstörungen könnte präventiv ein Feststellungsverfahren eingeleitet werden, damit die Schule zusätzliche sonderpädagogische Unterstützung erhält. Die Feststellung eines SPF könnte langfristig stigmatisierend wirken, da Kinder und Jugendliche einen SPF im Bildungsverlauf nur selten wieder verlieren (Pfahl, 2011; Powell, 2004; Wrase, 2015). Potenzielle negative Folgen auf Seiten der Lehrkräfte umfassen eine erhöhte Wahrnehmung für Auffälligkeiten und eine geringere Toleranz für heterogene Lernvoraussetzungen in der Lerngruppe (Wocken, 1996). Dadurch würden inklusive Bemü-

hungen nicht zu einer stärkeren Akzeptanz von Vielfalt, sondern paradoxerweise zu einer stärkeren Tendenz zur Diagnose und damit verbundenen Etikettierung von Abweichungen führen (Wocken, 1996).

Als Alternative schlagen verschiedene Forscherinnen und Forscher vor, sonderpädagogische Ressourcen für die Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* lerngruppenbasiert und nicht personengebunden bereitzustellen (Füssel & Kretschmann, 1993; Kottmann, 2006; Wocken, 1996; Wrase, 2015). Eine pauschale Mittelzuweisung kann beispielsweise an soziale Indikatoren des Schulbezirks gekoppelt werden (Rauer & Schuck, 2007; Wrase, 2015). Dadurch wird ein sonderpädagogischer Förderbedarf als Merkmal der Lerngruppe konzeptualisiert; individuelle Bedarfe in der Lerngruppe sollen anschließend flexibel und bedarfsgerecht bedient werden. Dieses Vorgehen kommt in manchen Bundesländern (z. B. Brandenburg, Hamburg) in Rahmen von Pilotprojekten schon zum Einsatz (Schröder-Lenzen & Krückels, 2015; Rauer & Schuck, 2007). Durch eine pauschale Ressourcenzuweisung pro Lerngruppe könnten Stigmatisierungseffekte reduziert werden (Wrase, 2015). Außerdem stehen sonderpädagogische Ressourcen der gesamten Lerngruppe zur Verfügung und könnten auch präventiv und kurzfristig eingesetzt werden (Wrase, 2015). Allerdings muss bei einer lerngruppenbezogenen Ressourcenzuweisung in geeigneter Weise sichergestellt werden, dass die Ressourcen bei den Kindern ankommen, die (sonderpädagogische) Unterstützung benötigen (Wrase, 2015).

2.3.3.3 Situation von Schülerinnen und Schülern mit SPF im internationalen Vergleich

Im internationalen Vergleich sonderpädagogischer Fördersysteme lassen sich drei Organisationsmodelle unterscheiden (Powell & Pfahl, 2012): Staaten mit dichotomen Modellen (z. B. Deutschland), in denen Schülerinnen und Schüler mit SPF überwiegend in separaten Förderschulen oder in allgemeinen Schulen unterrichtet werden; Staaten mit Kontinuumsmodellen (z. B. USA), in denen eine Vielzahl von Lernumgebungen für Kinder und Jugendliche mit SPF angeboten werden, die im Einzelfall flexibel angepasst werden können (z. B. Unterricht in getrennten Lerngruppen für bestimmte Schulfächer); und Staaten mit inklusiven Organisationsmodellen (z. B. Italien, Norwegen), in denen die meisten Kinder und Jugendlichen mit SPF in allgemeinen Schulen gemeinsam mit Schülerinnen und Schüler ohne SPF unterrichtet werden (Powell & Pfahl, 2012). Diese Organisationsmodelle sonderpädagogischer Förderung hängen mit strukturellen Merkmalen der Bildungssysteme zusammen (Biermann & Powell,

2014). So spiegelt sich Deutschlands hierarchisches und differenzierendes Bildungssystem auch im Umgang mit Schülerinnen und Schülern mit SPF wider. Im Gegensatz dazu implementieren Länder mit einem Gesamtschulsystem (Island, Schweden) eher ein inklusiveres sonderpädagogisches Fördersystem, in dem Kinder und Jugendliche mit SPF wohnortnah und für einen langen Zeitraum gemeinsam mit Schülerinnen und Schülern ohne SPF unterrichtet werden (Biermann & Powell, 2014). Ein separates sonderpädagogisches Fördersystem geht auch mit einer stärkeren Verknüpfung der Feststellung eines SPF und der Festlegung eines Förderortes einher (Powell & Pfahl, 2012). Während in Deutschland die Feststellung eines SPF in den Bereichen *Lernen* und *Sprache* immer noch mehrheitlich zu einer Einschulung oder Überweisung auf eine separate Förderschule führt, ist das in Ländern mit einem Gesamtschulsystem deutlich seltener der Fall (Biermann & Powell, 2014).

Im europäischen Vergleich hat Deutschland weder das exklusivste noch das inklusivste Bildungssystem (Powell & Pfahl, 2012). Vergleicht man zunächst die Förderquoten der europäischen Länder, so nimmt Deutschland mit einer Förderquote von circa 7 Prozent eine mittlere Position ein (Powell & Pfahl, 2012). Die Förderquoten variieren zwischen 1.5 Prozent in Italien und 18 Prozent in Finnland (Powell & Pfahl, 2012). Starke Unterschiede zwischen den Ländern gibt es auch in der Förderschulquote: In manchen Ländern werden nahezu alle Kinder und Jugendlichen mit SPF in separaten Förderschulen unterrichtet (z. B. Belgien), in anderen Ländern (z. B. Italien und Schweden) werden fast alle Schülerinnen und Schülern mit SPF in allgemeinen Schulen unterrichtet. So beträgt die Förderschulbesuchsquote in Italien und Schweden unter 0.1 Prozent, während sie in Deutschland bei 4.6 Prozent liegt. Damit liegt Deutschland über dem europäischen Durchschnitt von 2 Prozent (European Agency for Development in Special Needs Education, 2012). Beim internationalen Vergleich amtlicher Schulstatistiken gelten dieselben Einschränkungen und Herausforderungen wie beim Vergleich zwischen den Bundesländern in Deutschland. In Schweden wird zum Beispiel auf die offizielle Feststellung eines SPF verzichtet, so dass integrativ beziehungsweise inklusiv beschulte Kinder nicht in derselben Weise wie in Deutschland statistisch erfasst werden können (Biermann & Powell, 2014).

Zusammenfassend lässt sich das sonderpädagogische Fördersystem in Deutschland durch eine hohe Differenzierung und eine starke Rolle der Förderschulen beschreiben. Auch wenn Kinder und Jugendliche mit SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* zunehmend in allgemeinen Schulen unterrichtet werden, besucht die Mehrheit dieser Schülergruppe mit SPF immer noch Förderschulen. Dabei stellt sich die

Frage, wie die Beschulung in Förderschulen oder allgemeinen Schulen mit der schulischen Entwicklung von Kindern und Jugendlichen mit SPF zusammenhängt. Zur Beantwortung dieser Frage muss zunächst geprüft werden, wie Merkmale der schulischen Entwicklung dieser Schülergruppe valide erfasst werden können und welche methodischen Herausforderungen beim Einbezug dieser Schülergruppe in Schulleistungsstudien bestehen.

2.4 Methodische Herausforderungen beim Einbezug von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments

Grundlage der Dissertation sind Daten des IQB Ländervergleichs 2011, einem Large-Scale Assessment (LSA) in der Primarstufe. Ziel von LSAs ist es, festzustellen, inwieweit Kinder und Jugendliche Bildungsziele, wie beispielsweise Bildungsstandards, erreichen. LSAs identifizieren dabei Unterschiede im Erreichen der Bildungsstandards zwischen verschiedenen Bildungssystemen und zwischen verschiedenen Schülergruppen. Dadurch können sie Hinweise darauf geben, in welchen Bereichen Optimierungsbedarf besteht (Chudowsky & Pellegrino, 2003). Dieses Verständnis von LSAs impliziert, dass deren Ergebnisse alle Schülerinnen und Schüler eines Bildungssystems adäquat erfassen (Olson & Goldstein, 1997).

Die schulischen Kompetenzen von Schülerinnen und Schülern mit SPF wurden in LSAs lange Zeit nicht separat ausgewertet und berichtet, unter anderem auch aus dem Grund, dass diese Schülergruppe mehrheitlich in Förderschulen mit eigenen Lehrplänen beziehungsweise zielfferent in allgemeinen Schulen unterrichtet und deshalb nicht als Teil der Zielpopulation von LSAs definiert wurde (Heydrich et al., 2013; Müller, Sälzer, Mang, Heine et al., 2014; Thurlow, 2010). In den letzten Jahren stieg jedoch das Forschungsinteresse an den schulischen Kompetenzen der Schülergruppe mit SPF. Das liegt zum einen daran, dass der Anteil an Schülerinnen und Schülern mit SPF in den letzten 15 Jahren kontinuierlich angestiegen ist (s. Kapitel 2.3.3). Zum anderen werden Kinder und Jugendliche mit SPF immer häufiger an allgemeinen Schulen unterrichtet (s. Kapitel 2.3.3). Allerdings wurden die in LSAs eingesetzten Testinstrumente und Fragebögen nicht für Kinder und Jugendliche mit SPF entwickelt und erprobt. Daraus ergibt sich die Frage, ob die eingesetzten Testinstrumente dazu geeignet sind, die schulischen Kompetenzen dieser Schülergruppe vergleichbar zur Schülergruppe ohne SPF zu erfassen. Darüber hinaus muss für einen Schulartvergleich sichergestellt werden, dass die eingesetzten Testinstrumente die schulischen Kompetenzen und motivationalen Merkmale für

Schülerinnen und Schüler mit SPF in Förderschulen und allgemeinen Schulen in vergleichbarer Art und Weise erfassen. In diesem Kapitel werden Herausforderungen und mögliche Lösungsansätze für die Kompetenztestung von Kindern und Jugendlichen mit SPF beschrieben. Das Kapitel schließt mit einer Zusammenfassung des Forschungsstands zur Testgüte von Kompetenztests für Schülerinnen und Schüler mit SPF in Schulleistungstudien. Zwar liegt der Fokus in diesem Kapitel auf der Messung schulischer Kompetenzen, jedoch lassen sich die methodischen Herausforderungen auch auf die Erfassung motivationaler oder psychosozialer Merkmale übertragen (Nusser, Carstensen & Artelt, 2015; Schwinger et al., 2015).

2.4.1 Kompetenztestung bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments

In letzter Zeit wird vermehrt gefordert, Kinder und Jugendliche mit SPF in nationalen und internationalen Schulleistungstudien bei der Ergebnisauswertung und Darstellung der erreichten Kompetenzstände stärker zu berücksichtigen (National Center for Educational Statistics, 2010). Im Rahmen internationaler und nationaler LSAs werden Schülerinnen und Schüler mit SPF zwar in die Stichprobenziehung einbezogen, ihr Abschneiden in den standardisierten Leistungstests wird jedoch zumeist nicht separat ausgewertet und dargestellt (Heydrich et al., 2013; Müller, Sälzer, Mang, Heine et al., 2014; Thurlow, 2010). Diese Nichtberücksichtigung der Schülergruppe mit SPF in der Ergebnisdarstellung wirft die Frage auf, inwieweit die Ergebnisse aus LSAs und die daraus abgeleiteten Schlussfolgerungen auch für Schülerinnen und Schüler mit SPF aussagekräftig sind.

In Deutschland werden im Zuge der Umsetzung der UN-BRK (2008) zunehmend Bemühungen unternommen, die schulischen Kompetenzen, aber auch motivationale und psychosoziale Merkmale, von Kindern und Jugendlichen mit SPF im Rahmen von LSAs zu evaluieren. So wird in der UN-BRK (2008) darauf verwiesen, dass die Entwicklung eines inklusiveren Bildungssystems durch empirische Daten überprüft werden sollte (Artikel 31, UN-BRK, 2008). Nur wenn Schülerinnen und Schüler mit SPF in Schulleistungstudien berücksichtigt werden, können politische Entscheidungsträger aus deren Ergebnissen Wissen zur Umsetzung der UN-BRK (2008) ableiten. Zudem ist es wichtig, Merkmale der Lernumgebung von Kindern und Jugendlichen mit SPF und deren Auswirkungen auf schulische Outcomes zu evaluieren (Thurlow, 2002; Wrase, 2015).

Die Kompetenztestung von Schülerinnen und Schülern mit SPF in LSAs wird allerdings kontrovers diskutiert (Smith & Douglas, 2014; Thurlow, 2002). Kritiker einer stärkeren Berücksichtigung von Kindern und Jugendlichen mit SPF in LSAs verweisen auf der einen Seite darauf, dass LSAs in erster Linie schulische Leistungen auf der Systemebene evaluieren. LSAs zielen somit nicht darauf ab, Impulse für die Unterrichtsentwicklung und die konkrete Ausgestaltung von Inklusion bereitzustellen. Jedoch könnten die für LSAs konzipierten Tests auch für diese Zwecke verwendet werden. Dies würde zu einer starken Betonung schulischer Leistungen bei der Evaluation inklusiver Bemühungen führen (Smith & Douglas, 2014). Es wird argumentiert, dass auch solche Merkmale wie die sozialen Kompetenzen, die schulische Motivation und das Wohlbefinden der Schülerinnen und Schüler zur Beurteilung inklusiver Bemühungen herangezogen werden sollten (Smith & Douglas, 2014). Insbesondere die Förderung sozialer Kompetenzen und der Umgang mit Vielfalt sind explizite Ziele eines inklusiven Unterrichts und sollten deshalb auch bei dessen Evaluation berücksichtigt werden (Prenzel, 2013).

Darüber hinaus wird die Evaluation schulischer Inklusionsbemühungen auf Grundlage von LSAs dafür kritisiert, dass es aufgrund der überwiegend querschnittlichen Ausrichtung der Studien nicht möglich ist, Leistungsentwicklungen von Schülerinnen und Schülern mit SPF abzubilden. Längsschnittliche Studien seien besser zur Evaluation schulischer Inklusionsbemühungen geeignet, da sie auf Leistungszuwächse fokussieren und unterschiedliche Lernvoraussetzungen von Kindern und Jugendlichen mit SPF berücksichtigen (Morgan, Farkas & Wu, 2011; Wild et al., 2015). Schließlich wird auch argumentiert, dass die Umsetzung von Inklusion normativ festgelegt wird und keines Nachweises einer empirischen Wirksamkeit bedürfe (Emmerich, 2016; Veber, 2016). Demnach ermöglichen Studien mit Fokus auf der Unterrichtsebene eher Rückschlüsse darüber, wie ein inklusiver Unterricht gelingen kann. Im Gegensatz dazu würden LSAs den Blick zu stark darauf richten, ob bestimmte Aspekte von Inklusion, wie zum Beispiel die Beschulungsart, einen positiven Einfluss auf Kinder und Jugendliche mit SPF ausüben (Biewer, 2012; Smith & Douglas, 2014).

Auf der anderen Seite stellen LSAs eine wichtige Grundlage dar, um den Bildungserfolg der Schülergruppe mit SPF und ihre Bildungschancen für die weitere schulische und berufliche Ausbildung zu evaluieren. Außerdem könnte die Beschreibung der schulischen Kompetenzen und der Lernumgebung von Kindern und Jugendlichen mit SPF im Rahmen von LSAs dazu führen, dass diese Schülergruppe ebenfalls von daraus resultierenden bildungspolitischen Reformen profitiert (Bos et al., 2010; Thurlow, Ysseldyke, Gutman & Geenen, 1998). Eine

differenzierte Ergebnisdarstellung der schulischen Kompetenzen von Kindern und Jugendlichen mit SPF in LSAs ist darüber hinaus wichtig für eine adäquate Rückmeldung über die Leistungsfähigkeit des gesamten Bildungssystems (NAEP, 2010; Olson & Goldstein, 1997).

Zudem könnte der Einbezug von Schülerinnen und Schülern mit SPF in LSAs dazu beitragen, dass sie stärker als fähige Lerner anerkannt werden. Ihre stärkere Berücksichtigung in LSAs trage dazu bei, die Leistungserwartungen der Lehrkräfte gegenüber dieser Schülergruppe zu erhöhen (Thurlow, 2002). Schließlich besteht ein hoher gesellschaftlicher Konsens darüber, dass Schülerinnen und Schülern mit SPF Zugang zu den gleichen schulischen Erfahrungen ermöglicht werden sollte, die auch Kinder und Jugendliche ohne SPF erfahren (Pitoniak & Royer, 2001).

Zusammengefasst ersetzen LSAs somit zwar nicht die Erfassung individueller Lernfortschritte in akademischen, motivationalen und sozialen Bereichen, sie können aber ein nützliches Instrument zur Evaluation der schulischen Kompetenzen von Kindern und Jugendlichen mit SPF darstellen. Zum einen können sie Hinweise darauf geben, inwieweit auch die Schülergruppe mit SPF zentrale Bildungsziele erreicht, zum anderen haben sie aufgrund der großen Stichprobe und umfangreichen Test- und Befragungsinstrumente ein hohes Analysepotenzial.

2.4.2 Herausforderungen bei der Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

Eine zentrale Herausforderung für die Durchführung und Auswertung von Schulleistungstudien besteht darin, eine Balance zwischen den steigenden Teilnahmequoten von Schülerinnen und Schülern mit SPF auf der einen Seite und der Gewährleistung einer validen und reliablen Testung für alle Schülerinnen und Schüler auf der anderen Seite zu finden (Olson & Goldstein, 1997). Da die Testaufgaben in LSAs in der Regel nicht für die Schülergruppe mit SPF entwickelt und erprobt worden sind, kann nicht vorausgesetzt werden, dass sie für Schülerinnen und Schüler mit SPF dieselben Gütekriterien erfüllen beziehungsweise dasselbe Konstrukt abbilden wie für die Schülergruppe ohne SPF. Dies ist jedoch notwendig, um eine gemeinsame Skala für beide Schülergruppen definieren und ihre Leistungen darauf abbilden zu können. Damit verbunden ist die Frage der Vergleichbarkeit der Testergebnisse von Kindern und Jugendlichen mit SPF und Schülerinnen und Schülern ohne SPF (Lane & Leventhal, 2015). Eine eingeschränkte Vergleichbarkeit der Testergebnisse könnte dazu

führen, dass Unterschiede zwischen diesen beiden Schülergruppen im interessierenden Kompetenzkonstrukt durch die Testwerte nicht oder nur verzerrt abgebildet werden. Das betrifft sowohl den Vergleich zwischen Kindern und Jugendlichen mit SPF und Kindern und Jugendlichen ohne SPF, aber auch den Vergleich der schulischen Kompetenzen innerhalb der Schülergruppe mit SPF in Förderschulen und allgemeine Schulen. Für die vorliegende Arbeit ist letzteres besonders relevant, da ausgeschlossen werden muss, dass Unterschiede zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen auf unterschiedliche Messeigenschaften der Test- und Befragungsinstrumente zurückzuführen sind.

Vergleichbare psychometrische Eigenschaften der Kompetenztests zwischen Schülergruppen sind eine Voraussetzung für die valide Interpretation von Unterschieden in den Testwerten (Lane & Leventhal, 2015). Um Gruppenunterschiede in Testwerten valide interpretieren zu können, müssen strukturelle Aspekte der Testverfahren zwischen Schülergruppen vergleichbar sein (Shepard, 1993). Insbesondere sollten die interne Struktur der Leistungstests (z. B. die Anzahl zugrundeliegender Faktoren), die psychometrischen Eigenschaften der Testitems und die Schwierigkeiten der Items unter Kontrolle von Fähigkeitsunterschieden zwischen den Gruppen hinreichend ähnlich sein (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014). Darüber hinaus sollten die Korrelationen zwischen den Testwerten mit anderen Konstrukten zwischen den betrachteten Schülergruppen ähnlich ausfallen. Die theoriegeleitete Überprüfung von Zusammenhängen der Testergebnisse mit externen Kriterien, wie zum Beispiel mit den Testwerten anderer Leistungstests, stellt einen wichtigen Aspekt der Konstruktvalidität eines Testverfahrens dar (Hartig, Frey & Jude, 2012).

Die Interpretation der Testergebnisse von Schülerinnen und Schülern mit SPF kann durch mehrere Aspekte eingeschränkt sein, von denen drei im Folgenden näher betrachtet werden. Erstens können die vergleichsweise geringeren Leistungen von Schülerinnen und Schülern mit SPF dazu führen, dass die Leistungstests eine geringere Reliabilität aufweisen. Die Reliabilität von Leistungstests, die auf Modellen der Item-Response-Theory basieren, fällt im mittleren Bereich der Fähigkeitsverteilung am höchsten aus und ist an den Rändern der Fähigkeitsverteilung niedriger (Embretson & Reise, 2000). Für Schülerinnen und Schüler mit SPF ist im Vergleich zu Schülerinnen und Schülern ohne SPF eine deutlich geringere mittlere Fähigkeit (Abedi et al., 2011) sowie eine eingeschränkte Varianz in den Testleistungen zu erwarten (Lane & Leventhal, 2015). Dies kann die Messgenauigkeit der Leistungstests für Schülerinnen und Schüler mit SPF reduzieren (Elliott, Beddow, Kurz & Kettler, 2011).

Zweitens kann eine vergleichbare Interpretation der Testwerte von Kindern und Jugendlichen mit und ohne SPF durch konstruktferne schwierigkeitsgenerierende Merkmale der Testaufgaben (z. B. sprachliche Anforderungen, Antwortformate) eingeschränkt sein, insofern diese zwischen den Schülergruppen mit und ohne SPF unterschiedlich mit der Testbearbeitung zusammenhängen (Abedi et al., 2011; Südkamp, Pohl, Hardt, Jordan & Duchhardt, 2015). So könnten Schülerinnen und Schüler mit SPF zum Beispiel mit Multiple-Choice-Antwortformaten weniger vertraut sein als Schülerinnen und Schüler ohne SPF, wodurch solche Aufgaben für die Schülergruppe mit SPF differenziell schwieriger wären (Händel, Lockl, Heydrich, Weinert & Artelt, 2015). Dabei ist das Antwortformat ein konstruktfernes Merkmal, da die Testaufgaben nicht die Vertrautheit mit bestimmten Antwortformaten erfassen sollen. Außerdem könnten Kinder und Jugendliche mit SPF aufgrund geringerer Erfahrungen im Umgang mit standardisierten Leistungstests und aufgrund geringerer metakognitiver Fähigkeiten mehr Schwierigkeiten bei der Bearbeitung von bestimmten Aufgabenformaten haben als Schülerinnen und Schüler ohne SPF (Händel et al., 2015; Heydrich et al., 2013). Zur Analyse von Unterschieden in den Itemschwierigkeiten werden „differenzielle Itemfunktionen“ (*Differential Item Functioning*, DIF, Holland & Thayer, 1988) eingesetzt. Differenzielle Itemfunktionen liegen vor, wenn Schülerinnen und Schüler aus verschiedenen Gruppen trotz gleichem Fähigkeitsniveau eine spezifische Aufgabe mit unterschiedlich hoher Wahrscheinlichkeit lösen (Penfield & Camilli, 2007). Differenzielle Itemschwierigkeiten können dazu beitragen, dass die Testwerte der Schülerinnen und Schüler nur begrenzt miteinander vergleichbar sind (Abedi et al., 2011). Diese Einschränkung wird auch unter dem Aspekt der Testfairness diskutiert (Cole & Zieky, 2001).

Drittens kann die Testgüte bei der Erfassung schulischer Kompetenzen von Kindern und Jugendlichen mit SPF eingeschränkt sein, wenn sie mit den Testinhalten nicht ausreichend vertraut sind (Elliott et al., 2011). Unterschiede im Zugang zu testrelevanten Inhalten können in Abhängigkeit davon entstehen, ob Schülerinnen und Schüler mit SPF zielgleich oder zielfferent unterrichtet werden (Autorengruppe Bildungsberichterstattung, 2014). Zielfferent unterrichtete Schülerinnen und Schüler mit SPF werden nach speziellen Lehrplänen unterrichtet, die sich von den Lehrplänen für Schülerinnen und Schüler ohne SPF unterscheiden. Das trifft vor allem auf Schülerinnen und Schüler mit den Förderschwerpunkten *Lernen* und *geistige Entwicklung* zu (Autorengruppe Bildungsberichterstattung, 2016). Da sich die Kompetenztests in LSAs in der Regel an den Lehrplänen der Schülergruppe ohne SPF beziehungsweise den Bildungsstandards orientieren, könnten Schülerinnen und Schüler mit SPF insbesondere in Förderschulen mit den Anforderungen der Kompetenztests in LSAs

weniger vertraut sein. Dadurch könnte die Vergleichbarkeit der Testwerte zwischen zielgleich und zielfferent unterrichteten Schülerinnen und Schülern einschränken (Heydrich et al., 2013).

2.4.3 Umgang mit methodischen Herausforderungen bei der Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

Es stehen verschiedene Ansätze zur Verfügung, um den im vorherigen Kapitel beschriebenen Herausforderungen zu begegnen. Im Folgenden werden zwei Möglichkeiten vorgestellt: der Einsatz von Akkommodationen für die Schülergruppe mit SPF (z. B. Pitoniak & Royer, 2001) und die Erhöhung der Zugänglichkeit der Kompetenztests für alle Schülerinnen und Schüler (z. B. Beddow, 2011; Mace, 1991).

Akkommodationen stellen eine Möglichkeit dar, um die Testgüte von Kompetenztests für Schülerinnen und Schüler mit SPF zu erhöhen (Pohl, Südkamp, Hardt, Carstensen & Weinert, 2016; Südkamp, Pohl & Weinert, 2015). Akkommodationen sind Anpassungen in der Testgestaltung, die den Einfluss konstruktferner Merkmale auf die Testwerte der Schülerinnen und Schüler reduzieren (Pitoniak & Royer, 2001; Sireci, Scarpati & Li, 2005). Diese Anpassungen sollen den Einfluss spezifischer Lernschwierigkeiten und Beeinträchtigungen der Schülergruppe mit SPF auf die Testbearbeitung abschwächen (Elliott et al., 2011; Fuchs & Fuchs, 2001). Akkommodationen können das Präsentationsformat (z. B. Schriftgröße, Vorlesen der Testaufgaben), die zeitlichen Rahmenbedingungen (z. B. längere Bearbeitungszeit, Verteilung des Tests auf mehrere Testtage), das Antwortformat (z. B. Antwort im Testheft statt auf separatem Antwortbogen) und die Testumgebung (z. B. Einzeltestung) betreffen (Chiu & Pearson, 1999; Elliott et al., 2011; Sireci, 2004). Des Weiteren können auch Anpassungen an den Testaufgaben selbst vorgenommen werden.

Für Schülerinnen und Schüler mit SPF werden typischerweise Akkommodationen eingesetzt, bei der (1) die Testlänge reduziert beziehungsweise die zur Verfügung stehende Bearbeitungszeit pro Aufgabe erhöht wird und/oder (2) leichtere Testaufgaben eingesetzt werden (Cormier, Altman, Shyyan & Thurlow, 2010; Heydrich et al., 2013; Südkamp, Pohl & Weinert, 2015). Ziel der Anpassung der Testlänge ist es, Aufmerksamkeitsdefiziten beziehungsweise Ermüdungserscheinungen von Schülerinnen und Schülern mit SPF entgegenzuwirken (Müller, Sälzer, Mang, Gebhardt & Prenzel, 2013). Die Bearbeitung von Kompetenz-

tests erfordert ein ausreichendes Maß an Konzentration und Aufmerksamkeit über einen längeren Zeitraum. Durch eine reduzierte Testlänge können Unterschiede in der Aufmerksamkeit und Konzentrationsfähigkeit zwischen der Schülergruppe mit SPF angemessen berücksichtigt werden (Cormier et al., 2010; Heydrich et al., 2013).

Der Einsatz leichter Testaufgaben soll zum einen Frustration bei Schülerinnen und Schülern mit SPF aufgrund zu schwieriger Aufgaben verringern und deren Testmotivation erhöhen. So werden in manchen Schulleistungsstudien Aufgaben für jüngere Schülerinnen und Schüler eingesetzt (*Out-of-Level Testing*), um Schülerinnen und Schüler mit SPF nicht zu überfordern (Müller, Sälzer, Mang, Gebhardt et al., 2013; Müller, Sälzer, Mang, Heine et al., 2014; Südkamp, Pohl, Hardt et al., 2015). Zum anderen könnte dadurch die Präzision der Messung für Kinder und Jugendliche mit SPF gesteigert werden (Müller, Sälzer, Mang, Heine et al., 2014; Südkamp, Pohl, Hardt et al., 2015). Durch den Einsatz leichter Aufgaben kann eine höhere Passung der Itemschwierigkeiten und Personenfähigkeiten erreicht werden, die neben einer Erhöhung der Reliabilität auch einen positiven Einfluss auf die Testmotivation der Schülerinnen und Schüler mit SPF ausüben kann (Lane & Leventhal, 2015).

Eine Herausforderung beim Einsatz von Akkommodationen besteht darin, dass sie mit einer veränderten Testgestaltung einhergehen. Dabei muss sichergestellt werden, dass die Testwerte zwischen den Schülerinnen und Schülern, die Akkommodationen erhalten und denen, die keine Akkommodationen erhalten, vergleichbar sind. Akkommodationen können unter anderem zur Unterrepräsentation des zugrundeliegenden Konstruktes führen. So könnte der Verzicht auf schwierige Aufgaben in den Kompetenztests von Schülerinnen und Schülern mit SPF dazu beitragen, dass relevante Fähigkeiten nur in Ausschnitten und nicht in ihrer gesamten Breite erfasst werden (Tindal & Anderson, 2011). Zudem ist es schwer zu bestimmen, was konstruktferne Merkmale sind und ob stark vereinfachte Aufgaben noch dasselbe Konstrukt erfassen (Kettler & Elliott, 2010). Insbesondere bei Kindern und Jugendlichen mit den Förderschwerpunkten *Lernen* und *Sprache* hängen Lernschwierigkeiten häufig bedeutsam mit den zu messenden Konstrukten in Kompetenztests zusammen. Im Gegensatz dazu könnten zum Beispiel für Schülerinnen und Schüler mit dem Förderschwerpunkt *körperliche und motorische Entwicklung*, deren feinmotorische Fähigkeiten eingeschränkt sind, längere Bearbeitungszeiten und größere Antwortbögen bereitgestellt werden, ohne dass diese Akkommodationen direkt mit den schulischen Kompetenzen zusammenhängen. Vielmehr würden es solche Anpassungen der Schülergruppe erleichtern, die Aufgaben in den Kompetenztests zu beantworten. Schließlich liegt eine weitere Herausforderung darin, zu bestimmen,

welche Schülerinnen und Schüler angepasste Testaufgaben erhalten sollen (Lane & Leventhal, 2015).

Ein zweiter Ansatz, um die Testzugänglichkeit für Schülerinnen und Schüler mit SPF zu erhöhen besteht darin, konstruktferne Merkmale bei der Gestaltung der Testaufgaben für alle Schülerinnen und Schüler zu reduzieren und somit deren Zugänglichkeit insgesamt zu erhöhen (Beddow, 2011; Mace, 1991). Im Unterschied zu Akkommodation werden die Aufgaben bei diesem Ansatz für alle Schülerinnen und Schüler angepasst und in der gleichen Form vorgelegt. Eine höhere Zugänglichkeit der Aufgaben sollte dabei für alle Schülerinnen und Schüler von Vorteil sein, wobei bestimmte Schülergruppen bei der Testbearbeitung besonders profitieren können. Als theoretische Grundlage für diesen Ansatz dienen das *Universal Design* (Mace, 1991; Story, Mueller & Mace, 1998) und die *Accessibility Theory* (Beddow, 2011). Dabei werden zunächst allgemeine Richtlinien für die Aufgabenentwicklung entwickelt. Zu diesen Richtlinien zählen neben der Reduktion sprachlicher Komplexität auch klare Instruktionen sowie eine hohe Lesbarkeit der Aufgaben (Beddow, 2011). Ein sorgfältiger und elaborierter Prozess der Aufgabenkonstruktion kann den Bedarf an zusätzlichen Akkommodationen für bestimmte Schülergruppen verringern. Dadurch werden auch Vergleiche der Testwerte zwischen Schülergruppen erleichtert, da alle Schülerinnen und Schüler die gleichen Testaufgaben bearbeiten und gegebenenfalls unter den gleichen Bedingungen getestet werden können.

2.4.4 Forschungsstand zur Kompetenztestung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale Assessments

Im angloamerikanischen Sprachraum wurden zahlreiche Studien zur Vergleichbarkeit der Testwerte von Schülerinnen und Schülern mit SPF mit der Schülergruppe ohne SPF und zum Einfluss von Akkommodationen auf die Interpretation der Testwerte durchgeführt (Abedi et al., 2011; Harrison, Bunford, Evans & Owens, 2013; Lane & Leventhal, 2015; Smith & Douglas, 2014; Rogers, Christian & Thurlow, 2012). Die Ergebnisse dieser Studien, die sich mit den psychometrischen Eigenschaften und Aspekten der Vergleichbarkeit von Leistungstests für Schülerinnen und Schüler mit und ohne SPF befassen, sind uneinheitlich (s. für eine Zusammenfassung Lane & Leventhal, 2015). Es gibt Hinweise darauf, dass die faktorielle Struktur der Leistungstests in LSAs für Schülerinnen und Schüler mit und ohne SPF ähnlich

ist und nicht durch Akkommodationen beeinflusst wird (Cook, Eignor, Steinberg, Sawaki & Cline, 2009; Huynh, Meyer & Gallant, 2004). Weiterhin scheinen konstruktferne Merkmale das Abschneiden von Schülerinnen und Schülern mit SPF in Leseverständnistests nicht differenziell im Vergleich zur Schülergruppe ohne SPF zu beeinflussen (Kato, Moen & Thurlow, 2009). Diese Schlussfolgerungen lassen sich jedoch nicht auf Leistungstests im Allgemeinen übertragen. Je nach Jahrgangsstufe, betrachtetem Kompetenzbereich und Förderschwerpunkt unterscheiden sich die Befunde zu psychometrischen und strukturellen Eigenschaften der eingesetzten Leistungstests in Schulleistungsstudien (Lane & Leventhal, 2015). Außerdem limitieren geringe Stichprobenumfänge und geringe Überlappungsbereiche der Fähigkeitsverteilungen zwischen verschiedenen Schülergruppen die Ergebnisse zur Vergleichbarkeit von Testwerten (Sireci, 2009). Aus diesen Ergebnissen lässt sich ableiten, dass Aspekte der Testgüte und vergleichbaren Interpretation der Testwerte für jeden Leistungstest überprüft werden sollten.

Im Rahmen nationaler Erhebungen wurde jedoch bislang selten geprüft, ob sich die Testwerte der Schülergruppe mit SPF mit der Schülergruppe ohne SPF auf einer gemeinsamen Skala abbilden und vergleichend interpretieren lassen (Heydrich et al., 2013). Zwei Ausnahmen stellen Zusatzerhebungen im Rahmen von PISA 2012 (Müller, Sälzer, Mang, Heine et al., 2014) und des Nationalen Bildungspanels (NEPS; Heydrich et al., 2013; Südkamp, Pohl & Weinert, 2015) an Förderschulen dar. An die nationale Erhebung von PISA 2012 wurde ein Oversampling an Förderschulen mit den Förderschwerpunkten *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* angebunden (Müller, Sälzer, Mang, Heine et al., 2014). Dabei wurden 15-jährige Schülerinnen und Schüler mit SPF unter anderem in ihren schulischen Kompetenzen in Mathematik und im Leseverständnis getestet. Es wurden drei verschiedene Testbedingungen (reguläre PISA-Testhefte, einfachere Aufgaben, verkürzte Testhefte) im Hinblick auf die Vergleichbarkeit der Testwerte von Schülerinnen und Schülern mit und ohne SPF in allgemeinen Schulen geprüft. Dabei kamen Analysen zu fehlenden Werten, DIF-Analysen und Vergleiche der schulischen Kompetenzen zwischen den Testbedingungen zum Einsatz. Die Testbedingung mit verkürzten Testheften waren besser geeignet als die beiden anderen Testbedingungen, um die schulischen Kompetenzen der Schülerinnen und Schüler mit SPF valide und vergleichbar zur Schülergruppe ohne SPF abzubilden (Müller, Sälzer, Mang, Heine et al., 2014).

Südkamp, Pohl und Weinert (2015) untersuchten, inwieweit die Leseverständnistests in NEPS für Schülerinnen und Schüler an Förderschulen mit dem Förderschwerpunkt *Lernen* eine

reliable und zur Schülergruppe ohne SPF vergleichbare Messung ermöglichen. Außerdem wurde geprüft, inwieweit Akkommodationen die Testgüte für Schülerinnen und Schüler mit SPF erhöhen. Als Vergleichsgruppe wurden Schülerinnen und Schüler ohne SPF an Hauptschulen ausgewählt; beide Schülergruppen besuchten die fünfte Klasse. Die Testgüte wurde anhand des Anteils fehlender Werte, der Passung des Raschmodells, der Reliabilität und anhand von DIF-Analysen evaluiert. Im Standardtest ohne Akkommodationen zeigten sich höhere Anteile fehlender Werte bei Schülerinnen und Schülern mit dem Förderschwerpunkt *Lernen* als bei Schülerinnen und Schülern ohne SPF. Weiterhin wies ein bedeutsamer Anteil der Aufgaben für Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* eine schlechte Passung des Raschmodells auf. Außerdem war die Reliabilität des Leseverständnistests für diese Schülergruppe deutlich geringer als für Hauptschülerinnen und -schüler. Außerdem zeigte sich bei einem bedeutsamen Anteil der eingesetzten Aufgaben ein starker DIF. Insgesamt war die Vergleichbarkeit der Testwerte zwischen den betrachteten Schülergruppen somit stark eingeschränkt. Der Einsatz von Akkommodationen (reduzierte Testzeit bzw. einfachere Aufgaben) führte zu einer höheren Testgüte bei der Erfassung der schulischen Kompetenzen von Schülerinnen und Schülern mit dem Förderschwerpunkt *Lernen* (s. a. Südkamp, Pohl, Hardt et al., 2015).

Zusammenfassend zeigte sich in beiden Studien, dass angepasste Testbedingungen in Form von verkürzten Testheften besser geeignet scheint, die schulischen Kompetenzen der Schülerinnen und Schüler mit SPF zu erfassen, als reguläre Testbedingungen (Müller, Sälzer, Mang, Heine et al., 2014; Südkamp, Pohl & Weinert, 2015). Beide Studien bezogen sich dabei auf eine vergleichbare Messung der schulischen Kompetenzen von Schülerinnen und Schüler mit SPF in Förderschulen und Schülerinnen und Schülern ohne SPF. Im Rahmen von LSAs werden jedoch auch die Leistungen von Schülerinnen und Schülern mit SPF in allgemeinen Schulen erfasst. Zudem besteht erhebliches wissenschaftliches und bildungspolitisches Interesse an der Frage, inwieweit die Leistungsentwicklung von Schülerinnen und Schülern mit SPF durch die Art der Beschulung beeinflusst wird. Daher ist es wichtig zu prüfen, ob die eingesetzten Testverfahren Fähigkeitsunterschiede bei Kindern mit SPF in beiden Schularten gleichermaßen valide und reliabel abbilden können.

2.5 Der Zusammenhang zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

In diesem Kapitel werden theoretische Annahmen und empirische Befunde zum Zusammenhang zwischen der Beschulungsart und Merkmalen der schulischen Entwicklung von Kindern und Jugendlichen beschrieben. In Kapitel 2.5.1 wird ein Rahmenmodell zur Vorhersage schulischer Outcomes durch die schulische und familiäre Lernumgebung vorgestellt. Auf dieser Grundlage werden Schularten als differenzielle Lern- und Entwicklungsmilieus beschrieben und mögliche Erklärungen für Unterschiede in der schulischen Entwicklung in Abhängigkeit von der besuchten Schulart vorgestellt (Kapitel 2.5.2). Diese Annahmen zu Schularteffekten werden anschließend auf die Unterschiede in schulischen Leistungen (Kapitel 2.5.3) und auf die Merkmale der schulischen Motivation (Kapitel 2.5.4) von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen angewendet.

2.5.1 Rahmenmodell zum Zusammenhang von schulischer und familiärer Lernumgebung mit schulischen Outcomes

Die Vielfalt an Inklusionsdefinitionen führt dazu (s. Kapitel 2.2), dass sich Studien in ihrer Operationalisierung und empirischen Untersuchung von Inklusion unterscheiden (Forlin, Chambers, Loreman, Deppler & Sharma, 2013). Um verschiedene Untersuchungsansätze zu vergleichen, bietet sich das Kontext-Input-Prozess-Output-Modell (Scheerens, 1990, 2008) als konzeptueller Rahmen an (European Agency for Development in Special Needs Education, 2011; Gresch & Piezunka, 2015). Anhand dieses Modells können relevante Kontext- und Inputfaktoren sowie vermittelnde Prozesse identifiziert werden, die Merkmale der schulischen Entwicklung von Kinder und Jugendlichen beeinflussen (Scheerens, 2004, 2008).

In Abbildung 4 (nach Scheerens, 2008, S. 27 & S. 53) sind die vier Kernelemente (Kontext, Input, Prozess, Outcome) des Modells schematisch dargestellt. Zum *Kontext* gehören zeitlich vergleichsweise stabile Merkmale, welche Rahmenbedingungen des Unterrichts festlegen. Dazu zählen unter anderem Richtlinien zur Feststellung eines SPF und Differenzierung verschiedener Förderschwerpunkte (s. Kapitel 2.3.1 und 2.3.2), länderspezifische Ausgestaltungen sonderpädagogischer Förderung (s. Kapitel 2.3.2 und 2.3.3) und die Zuweisung von Ressourcen zu Schülerinnen und Schülern mit SPF. Zum *Input* zählen Rahmenbedingungen

auf Schüler-, Klassen-, und Schulebene, die in engerer Beziehung zu Lernprozessen stehen. Dazu gehört auch die Beschulung in Förderschulen oder allgemeinen Schulen, die mit schulartspezifischen Lehrplänen und Stundentafeln (s. Kapitel 2.5.2.3), Merkmalen der Schülerinnen und Schüler (z. B. Schweregrad der Beeinträchtigung, Vorwissen), Merkmalen der Klassenzusammensetzung (s. Kapitel 2.5.2.2) und Merkmalen der Lehrkräfte (z. B. Qualifikation und Erfahrung im Unterricht mit Schülerinnen und Schülern mit SPF, Einstellungen zum gemeinsamen Unterricht) einhergeht. Die *Prozessebene* lässt sich in Merkmale der schulischen Organisation (z. B. verbrachte Zeit im gemeinsamen Unterricht) und Merkmale der Unterrichtsgestaltung (z. B. Individualisierung im Unterricht, *Peer-Tutoring*, s. Kapitel 2.5.2.2) differenzieren. Die Ebene des *Outputs* beinhaltet sowohl kognitive als auch nicht-kognitive Erträge und Folgen von Bildung (z. B. schulische Leistungen und Abschlüsse, Motivation, schulisches Wohlbefinden, soziale Kompetenzen).

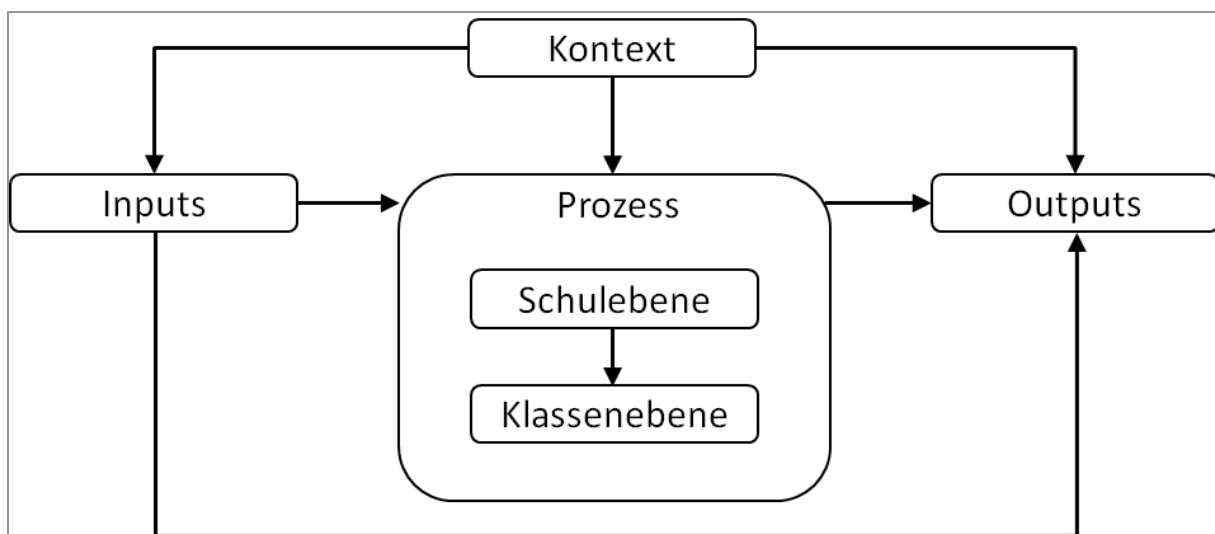


Abbildung 4. Kontext-Input-Prozess-Output-Modell zur Untersuchung von Effekten der Schule und des Unterrichts auf Schülerinnen und Schüler (nach Scheerens, 2008, S. 27 & S. 55).

In dieser Dissertation werden Zusammenhänge zwischen der Beschulungsart (Input) und den schulischen Leistungen sowie Merkmalen der schulischen Motivation (Output) von Kindern mit SPF betrachtet. Um eine genauere Schätzung der Schularteffekte zu erzielen, werden relevante individuelle Ausgangsbedingungen der Schülerinnen und Schüler mit SPF in Förderschulen und allgemeinen Schulen berücksichtigt (Scheerens, 2008). Außerdem wird in Einklang mit dem Kontext-Input-Prozess-Output-Modell ein Mehrebenenansatz verfolgt, bei

dem individuelle Merkmale der Schülerinnen und Schüler in Beziehung zur Klassenkomposition in Förderschulen und allgemeinen Schulen gesetzt werden.

2.5.2 Förderschulen und allgemeine Schulen als differenzielle Lern- und Entwicklungsmilieus

Eine zentrale Annahme des Modells von Scheerens (2008) ist, dass die schulische Entwicklung nicht nur von den individuellen Lernvoraussetzungen und der sozialen und kulturellen Herkunft der Schülerinnen und Schüler abhängt, sondern dass Merkmale der Lernumgebung zu unterschiedlichen Entwicklungsverläufen der Schülerinnen und Schüler beitragen. Diese Annahme wurde von der Forschergruppe um Baumert auf Unterschiede in der Lernumgebung zwischen verschiedenen Schularten übertragen (Baumert & Köller, 1998; Baumert, Köller & Schnabel, 2000; Baumert, Stanat & Watermann, 2006). Dabei gehen sie davon aus, dass Schularten differenzielle Lern- und Entwicklungsmilieus darstellen. Demnach unterscheidet sich die Lernumgebung zwischen Schularten systematisch, so dass Kinder und Jugendliche je nach besuchter Schulart unterschiedliche Entwicklungschancen erhalten (Baumert et al., 2006). Dies wirkt sich auf die Leistungsentwicklung von Kindern und Jugendlichen verschiedener Schularten aus. Differenzielle Entwicklungsverläufe treten auch nach Berücksichtigung individueller Lernvoraussetzungen auf, das heißt, Merkmale der Lernumgebung in verschiedenen Schularten wirken *zusätzlich* zu individuellen Lernvoraussetzungen auf die schulische Entwicklung von Kindern und Jugendlichen (Baumert et al., 2006). Differenzielle Lern- und Entwicklungsmilieus führen in erster Linie zu unterschiedlichen Leistungsentwicklungen, sie beeinflussen aber auch die motivationale und psychosoziale Entwicklung der Schülerinnen und Schüler (Baumert & Köller, 1998). Unterschiedliche Entwicklungsverläufe von Kindern und Jugendlichen in Abhängigkeit von der besuchten Schulart können durch mindestens drei Faktoren erklärt werden (Baumert et al., 2006, Dumont, Neumann, Maaz & Trautwein, 2013, Harris, 2010).

Erstens können *Unterschiede in den individuellen Lernausgangslagen* (z. B. kognitive Fähigkeiten, Vorwissen, familiäre Merkmale) dazu beitragen, dass sich schulische Leistungen und weitere Merkmale der schulischen Entwicklung von Kindern und Jugendlichen zwischen Schularten unterscheiden. Demnach könnten Schülerinnen und Schüler mit höherem Vorwissen oder höheren kognitiven Fähigkeiten auch größere Lernfortschritte als Kinder mit niedrigeren Ausgangsleistungen erzielen (Baumert, Nagy & Lehmann, 2012; Morgan, Farkas

& Hibel, 2008; Stanovich, 1986). Diese differenziellen Entwicklungsverläufe werden als individueller Matthäuseffekt (Stanovich, 1986) bezeichnet und tragen zu Schulartunterschieden bei. Allerdings handelt es sich dabei streng genommen nicht um einen Erklärungsansatz für Schularten als differenzielle Lern- und Entwicklungsmilieus. Bei der Annahme differenzieller Lern- und Entwicklungsmilieus geht man vielmehr davon aus, dass Merkmale der Lernumgebung unabhängig von und zusätzlich zu individuellen Lernvoraussetzungen zu unterschiedlichen Entwicklungsverläufen führen. Deshalb wird dieser Erklärungsansatz im Folgenden nicht weiter beschrieben.

Eine zweite Erklärung für differenzielle Entwicklungsverläufe von Kindern und Jugendlichen sind *Kompositionseffekte* (s. Kapitel 2.5.2.2; Harris, 2010). Darunter werden Auswirkungen der Schülerzusammensetzung einer Klasse oder Schule verstanden, die unabhängig von und zusätzlich zu individuellen Schülermerkmalen auf die schulische Entwicklung der Kinder und Jugendlichen wirken (Thrupp, Lauder & Robinson, 2002). Unterschiede in der Schülerzusammensetzung zwischen Schularten beziehen sich in erster Linie auf das Leistungs- beziehungsweise Fähigkeitsniveau. Allerdings kann sich die Schülerschaft auch in ihrer soziokulturellen Zusammensetzung unterscheiden, da zum Beispiel Merkmale der soziokulturellen Herkunft mit schulischen Leistungen kovariieren (z. B. Haag, Böhme & Stanat, 2012; Richter, Kuhl & Pant, 2012).

Eine dritte Erklärung für differenzielle Entwicklungsverläufe in Abhängigkeit von der besuchten Schulart sind *institutionelle Effekte* (s. Kapitel 2.5.2.3). Darunter werden unterschiedliche pädagogisch-didaktische Unterrichtstraditionen der Schularten wie zum Beispiel schulartspezifische Lehrpläne zusammengefasst, die die schulische Entwicklung von Kindern und Jugendlichen beeinflussen (Baumert et al., 2006; Maaz, Baumert & Trautwein, 2010; Wocken, 2005). Darüber hinaus können Wechselwirkungen zwischen der Schülerkomposition und institutionellen Merkmalen auftreten, etwa wenn Lehrkräfte das Anforderungsniveau ihres Unterrichts an die Leistungsstärke der Lerngruppe anpassen (Jerusalem, 1997). In den folgenden Kapiteln werden diese Erklärungen ausführlicher beschrieben. Dabei wird auch auf Interaktionen zwischen kompositionellen und institutionellen Effekten eingegangen (s. Kapitel 2.5.2.4).

2.5.2.1 Entstehung schulartspezifischer Schülerzusammensetzungen

Schülerinnen und Schüler werden primär auf Grundlage ihrer schulischen Leistungen (z. B. Schulnoten) und ihrer Bildungsgeschichte (z. B. Einschulungsalter, Klassenwiederholungen; Tillmann, 2006) zu unterschiedlichen Schularten zugewiesen. Durch diese Zuweisung unterscheiden sich Schularten in der fähigkeitsbezogenen, aber auch soziokulturellen Zusammensetzung ihrer Schülerschaft (Kornmann, 2006; Maaz et al., 2010). Zum einen sind Merkmale der sozialen Herkunft mit schulischen Leistungen der Schülerinnen und Schüler konfundiert (z. B. Ditton, Krüsken & Schauenberg, 2005; Haag et al., 2012; Jackson, Erikson, Goldthorpe & Yaish, 2007; Richter et al., 2012). Zum anderen scheinen Merkmale der sozialen und zuwanderungsbezogenen Herkunft zusätzlich zu schulischen Leistungen bei Einschulungs- oder Übergangsentscheidungen eine Rolle zu spielen (Auernheimer, 2010; Trautwein, Nagy & Maaz, 2011). Da keine einheitlichen Kriterien für die Förderschulüberweisung existieren (s. Kapitel 2.3.2), entsteht bei der Schulartzuweisung ein gewisser Entscheidungsspielraum für schulische Akteure und Erziehungsberechtigte (Kronig, 2001; Sälzer et al., 2015). Dadurch könnten auch herkunftsbezogene Merkmale der Schülerinnen und Schüler die Schulartzuweisung beeinflussen.

Sowohl Auswertungen amtlicher Schulstatistiken als auch empirische Studien zeigen, dass Kinder und Jugendliche in Förderschulen (v. a. im Förderschwerpunkt *Lernen*) über geringere sozioökonomische Ressourcen verfügen und häufiger aus bildungsfernen Familien stammen als Kinder und Jugendliche in allgemeinen Schulen (Autorengruppe Bildungsberichterstattung, 2012, 2014; Bos et al., 2010; Klein, 2001; Opp et al., 2008; Rauer & Schuck, 2007; Schumann, 2007; Wocken, 2000, 2005). So weisen Eltern von Förderschülerinnen und Förderschülern z. B. deutliche niedrigere Bildungsabschlüsse auf als Eltern von Kindern und Jugendlichen in allgemeinen Schulen; die Mehrzahl der Eltern der Kinder an Förderschulen verfügt höchstens über einen Hauptschulabschluss (Autorengruppe Bildungsberichterstattung, 2010; Wocken, 2000, 2005). Auch der Anteil der Eltern ohne einen berufsqualifizierenden Abschluss ist an Förderschulen im Vergleich zu sonstigen Schularten doppelt so hoch (Autorengruppe Bildungsberichterstattung, 2010; Wocken, 2000, 2005).

Des Weiteren ist ein Drittel der Eltern der Kinder und Jugendlichen an Förderschulen nicht erwerbstätig; in allen anderen betrachteten Schularten ist diese Quote deutlich niedriger (Autorengruppe Bildungsberichterstattung, 2010; Wocken, 2000, 2005). Wocken (2000, 2005) zeigte, dass Schülerinnen und Schüler an Förderschulen mit dem Förderschwerpunkt

Lernen über ein geringeres häusliches Anregungsniveau (u. a. häuslicher Buchbestand) verfügen als Schülerinnen und Schüler anderer Schularten. Zudem sind Kinder und Jugendliche mit einem Zuwanderungshintergrund an Förderschulen überrepräsentiert, wobei diese Unterschiede jedoch in Abhängigkeit des betrachteten Förderschwerpunktes und der betrachteten Herkunftsgruppe variieren (Brodesser, 2016; Dietze, 2011; Kemper & Weishaupt, 2011; Kornmann, 2006; Kronig, 2001, 2003; Powell & Wagner, 2002; Wagner & Powell, 2003).

Zur Erklärung dieser Überrepräsentation soziokulturell benachteiligter Schülergruppen an Förderschulen lassen sich individuumszentrierte, interaktionale und institutionelle Ansätze unterscheiden (Gomolla & Radtke, 2009; Kornmann, 2006; Kronig, 2003).

Individuumszentrierte Ansätze konzentrieren sich auf die Lernvoraussetzungen und Lebenslagen von Kindern und Jugendlichen aus zugewanderten und sozial schwachen Familien. Ein theoretisches Modell zur Entstehung sozialer Ungleichheiten im Bildungssystem wurde von Boudon (1974) beschrieben. Boudon (1974) geht davon aus, dass schulische Leistungen und Bildungsentscheidungen vom sozialen und zuwanderungsbezogenen Hintergrund der Familien beeinflusst werden. Wichtig ist dabei die Differenzierung zwischen primären und sekundären Effekten der sozialen Herkunft (vgl. Baumert & Schümer, 2001). Primäre Herkunftseffekte basieren auf dem sozio-ökonomischen und kulturellen Kapital der Familien und wirken sich auf die schulischen Leistungen der Schülerinnen und Schüler aus. Es wird davon ausgegangen, dass Kinder und Jugendliche aus sozial schwachen Familien auch deshalb niedrigere Schulleistungen erzielen, weil ihr häusliches Anregungsniveau und kulturelles Kapital geringer ausgeprägt ist (McElvany, Becker & Lüdtke, 2009).

Von sekundären Effekten spricht man, wenn Schulwahlentscheidungen herkunftsspezifisch getroffen werden. Dadurch erzielen Kinder und Jugendliche aus sozial schwachen Familien selbst bei vergleichbaren Schulleistungen niedrigere Bildungsabschlüsse. Sekundäre Effekte spiegeln spezifische Bildungsentscheidungen und Übergangsneigungen der Eltern wider. Bildungsentscheidungen werden unter anderem durch ökonomische Kosten-Nutzen-Abwägungen im Umgang mit Bildungsressourcen und den Bildungsaspirationen der Eltern beeinflusst (Stocké, 2013). Dazu gehören Einschätzungen zu voraussichtlichen Kosten (z. B. Investitionsaufwand, Opportunitätskosten) und Erträgen (z. B. berufliche Chancen und Einkommen) von unterschiedlichen Bildungsangeboten (Becker, 2007). So könnten Eltern aus sozial schwachen Schichten bei einer Förderschulempfehlung ihren Entscheidungsspielraum und die Konsequenzen einer solchen Überweisung für den weiteren Bildungsverlauf ihres Kindes unterschätzen (Schütz, West & Wößmann, 2007).

Interaktionale Erklärungsansätze gehen davon aus, dass die Förderschulüberweisung von Kindern und Jugendlichen aus sozial schwachen Elternhäusern oder zugewanderten Familien durch soziale Stereotype und Erwartungshaltungen der Lehrkräfte und anderer schulischer Akteure beeinflusst wird (Diamond, Randolph & Spillane, 2004; Kronig, 2003; Rist, 2000). Selektionsentscheidungen basieren auch auf Lehrkraftbeurteilungen des Sozialverhaltens und der Gesamtentwicklung der Schülerinnen und Schüler (Lentz & Radtke, 1994, s. Kapitel 2.3.2). Das Sozial- und Lernverhalten von Schülerinnen und Schülern mit einem Zuwanderungshintergrund oder aus sozial schwachen Schichten könnte aufgrund ihrer Herkunft eher als problematisch oder nicht altersgemäß eingeschätzt werden (Gomolla & Radtke, 2009; Morgan et al., 2015). So könnten Lehrkräfte die Fähigkeiten von Schülerinnen und Schülern mit geringeren sozialen Ressourcen eher unterschätzen (Kronig, 2003) oder bei Kindern und Jugendlichen mit einem Zuwanderungshintergrund eine geringere Passung zwischen den akademischen und sozialen Standards der Schule mit den familiären Leistungs- und Verhaltensstandards feststellen (Kronig, 2003, Morgan et al., 2015). Solche Mechanismen können dazu führen, dass Kinder aus sozial schwachen beziehungsweise zugewanderten Familien bei vergleichbar geringeren Leistungen mit einer höheren Wahrscheinlichkeit an eine Förderschule überwiesen werden.

Allerdings können differenzielle Bewertungsstandards auch zugunsten von Kindern und Jugendlichen aus sozial schwachen oder zugewanderten Familien ausfallen (Morgan et al., 2015). So wäre es möglich, dass Lehrkräfte niedrigere Leistungen von Schülerinnen und Schülern aus sozial schwachen oder zugewanderten Familien stärker auf geringe familiäre Ressourcen oder mangelnde Sprachkenntnisse zurückführen (Morgan et al., 2015). Darüber hinaus besuchen Schülerinnen und Schüler aus sozial schwachen Familien häufiger Kindertagesstätten oder Schulen mit einem höheren Anteil an leistungsschwächeren Kindern (Morgan et al., 2015). In einer solchen Lernumgebung dürften sich diese Kinder weniger stark von ihren Mitschülerinnen und Mitschülern unterscheiden, was dazu führen könnte, dass die Feststellung eines SPF seltener eingeleitet wird (*Froschteicheffekt*; Hibel, Farkas & Morgan, 2010).

Institutionelle Erklärungsansätze führen die Überrepräsentation von Schülerinnen und Schülern aus sozial schwachen und zugewanderten Familien in Förderschulen auf institutionelle Merkmale des Bildungssystems zurück (Kronig, 2003; Wagner & Powell, 2003). Institutionelle Merkmale umfassen Regelungen, Normen und Routinen auf Schulebene beziehungsweise auf Ebene der Schulverwaltung, die das Handeln schulischer Akteure

beeinflussen (Gomolla & Radtke, 2009; Lentz & Radtke, 1994). Dazu gehören auch ungeschriebene Regeln und Gewohnheiten, die meist nicht intentional ablaufen (Gomolla & Radtke, 2009). Institutionelle Erklärungsansätze und interaktionale Erklärungen überschneiden sich, da sie den Blick auf schulische Akteure anstatt auf die Schülerinnen und Schüler und ihr familiäres Umfeld richten. Ein zentraler Unterschied zwischen beiden Erklärungsansätzen liegt darin, dass institutionelle Mechanismen stärker auf strukturelle und organisatorische Bedingungen des Bildungssystems fokussieren, während interaktionale Erklärungen das Handeln der beteiligten pädagogischen Fachkräfte in den Vordergrund stellen. So kann beispielsweise die fehlende Überprüfung muttersprachlicher Fähigkeiten bei der Einleitung eines Feststellungsverfahrens durch die Klassenlehrkraft als ein interaktionaler Erklärungsansatz für eine mögliche Überrepräsentation gesehen werden, während die fehlende Verankerung und Bereitstellung angemessener muttersprachlicher Testverfahren beim Feststellungsverfahren eher eine institutionelle Erklärung darstellen würde (Lindsay, Pather & Strand, 2006; Morgan et al., 2015).

Eine zentrale These institutioneller Erklärungsansätze ist, dass der überproportional hohe Anteil an Kindern und Jugendlichen aus sozial schwachen beziehungsweise zugewanderten Familien an Förderschulen eine Form der institutionellen Diskriminierung darstellt (Lentz & Radtke, 1994). Demnach können organisatorische Rahmenbedingungen dazu führen, dass Schülerinnen und Schüler aus sozial schwachen oder zugewanderten Familien systematisch benachteiligt werden. So könnten institutionelle Faktoren, wie der Ausbau oder die Aufrechterhaltung von Förderschulen, vor dem Hintergrund sinkender Schülerzahlen zu einer vermehrten Förderschulüberweisung von Kindern und Jugendlichen aus sozial schwachen beziehungsweise zugewanderten Familien führen, da sie im Durchschnitt niedrigere schulische Leistungen aufweisen (Gomolla & Radtke, 2009).

Hinweise für diese Form von institutioneller Diskriminierung sind z. B. regionale Unterschiede im Förderschulangebot, die zu unterschiedlichen Förderschulüberweisungsraten in Abhängigkeit des sozialen Status beziehungsweise Zuwanderungshintergrundes führen können (Powell & Wagner, 2002). Als ein Beispiel führen Lentz und Radtke (1994) an, dass die Abschaffung der Vorbereitungsklassen in Nordrhein-Westfalen dazu beitrug, dass Schülerinnen und Schüler mit einem Zuwanderungshintergrund vermehrt in Förderschulen mit dem Schwerpunkt *Lernen* eingeschult wurden, um leistungshomogene Lerngruppen zu bilden. Von diesem Anstieg waren Kinder griechischer Herkunft aufgrund der Schaffung privater Ergänzungsschulen jedoch nicht betroffen (Lentz & Radtke, 1994). Diese herkunfts-

spezifischen Förderschulüberweisungen wurden darauf zurückgeführt, dass Förderschulen eine entlastende Funktion für allgemeine Schulen übernommen haben, die vormals von Vorbereitungsklassen und im Falle Kinder griechischer Herkunft von privaten Ergänzungsschulen eingenommen wurde (Lentz & Radtke, 1994).

2.5.2.2 Kompositionseffekte

Förderschulen und allgemeine Schulen unterscheiden sich in der leistungsbezogenen und soziokulturellen Zusammensetzung ihrer Schülerschaft. Die Schülerschaft in allgemeinen Schulen ist durchschnittlich leistungstärker und verfügt über mehr sozioökonomische Ressourcen als die Schülerschaft in Förderschulen (Brodesser, 2016; Klein, 2001; Wocken, 2005). Diese Unterschiede in der Zusammensetzung der Schülerschaft können in Form von Kompositionseffekten zu differenziellen Entwicklungsverläufen von Kindern und Jugendlichen in verschiedenen Schularten beitragen. Von Kompositionseffekte spricht man, wenn die Zusammensetzung der Schülerschaft zusätzlich zu und unabhängig von individuellen Merkmalen der Schülerinnen und Schüler die schulische Entwicklung beeinflusst (Harker & Tymms, 2004; Lüdtke, Robitzsch & Köller, 2002). Die Schülerzusammensetzung wird dabei meistens durch die Aggregation individueller Merkmale auf Klassen- oder Schulebene erfasst (Wilkinson, 2002).

Als zentrale Beschreibungsmerkmale der Schülerzusammensetzung gelten folgende Dimensionen (Baumert et al., 2006):

- a) das Fähigkeits- und Leistungsniveau einer Lerngruppe
- b) die soziokulturelle Zusammensetzung (z. B. erfasst über das kulturelle und soziale Kapital der Eltern)
- c) die ethnisch-kulturelle Zusammensetzung (z. B. erfasst über den Anteil an Schülerinnen und Schülern nichtdeutscher Herkunftssprache)
- d) die Häufigkeit sozialer Risikofaktoren (z. B. Anteil der Schülerinnen und Schüler in instabilen Familienverhältnissen)
- e) der Anteil an Schülerinnen und Schülern, die Belastungen oder Schwierigkeiten in ihrer schulischen Karriere erlebten (z. B. Anteil an Klassenwiederholerinnen und -wiederholern)

Diese Dimensionen ließen sich anhand konfirmatorischen Faktorenanalysen differenzieren (Baumert et al., 2006). Allerdings ist deren Bedeutsamkeit für die Vorhersage der schulischen Entwicklung von Kindern und Jugendlichen eine empirische Frage. So könnte in Schulen, in denen eine Konstellation von mehreren ungünstigen Kompositionsmerkmalen vorliegt, ein besonders ungünstiges Lern- und Entwicklungsmilieu entstehen.

Eine zentrale Annahme kompositioneller Effekte besteht darin, dass sich die Schülerzusammensetzung nicht direkt auf die individuelle schulische Entwicklung auswirkt, sondern dass sie über verschiedene Prozesse vermittelt wird. Ein erstes Vermittlungsmodell wurde von Dreeben und Barr (1988) aufgestellt, welches auf die individuelle Leistungsentwicklung fokussiert. Demnach beeinflusst die Schülerzusammensetzung sowohl die Interaktion zwischen Lehrkräften und der Lerngruppe (*Instruktionsprozesse*) als auch die Interaktionen zwischen Schülerinnen und Schülern (*Referenzgruppenprozesse*). Baumert et al. (2006) gehen in ihrem Vermittlungsmodell (Abbildung 5) davon aus, dass die eingangs beschriebenen fünf Dimensionen zur Beschreibung der Schülerzusammensetzung gemeinsam mit institutionellen Merkmalen der Schulart die individuelle schulische Entwicklung beeinflussen. Diese Einflüsse werden durch Interaktionen zwischen Lehrkräften, Schülerinnen und Schülern und Eltern vermittelt. Fasst man beide Modelle (Baumert et al., 2006; Dreeben & Barr, 1988) zusammen, lassen sich vier Vermittlungsprozesse unterscheiden, die sich wechselseitig beeinflussen (Hallinan, 1988).

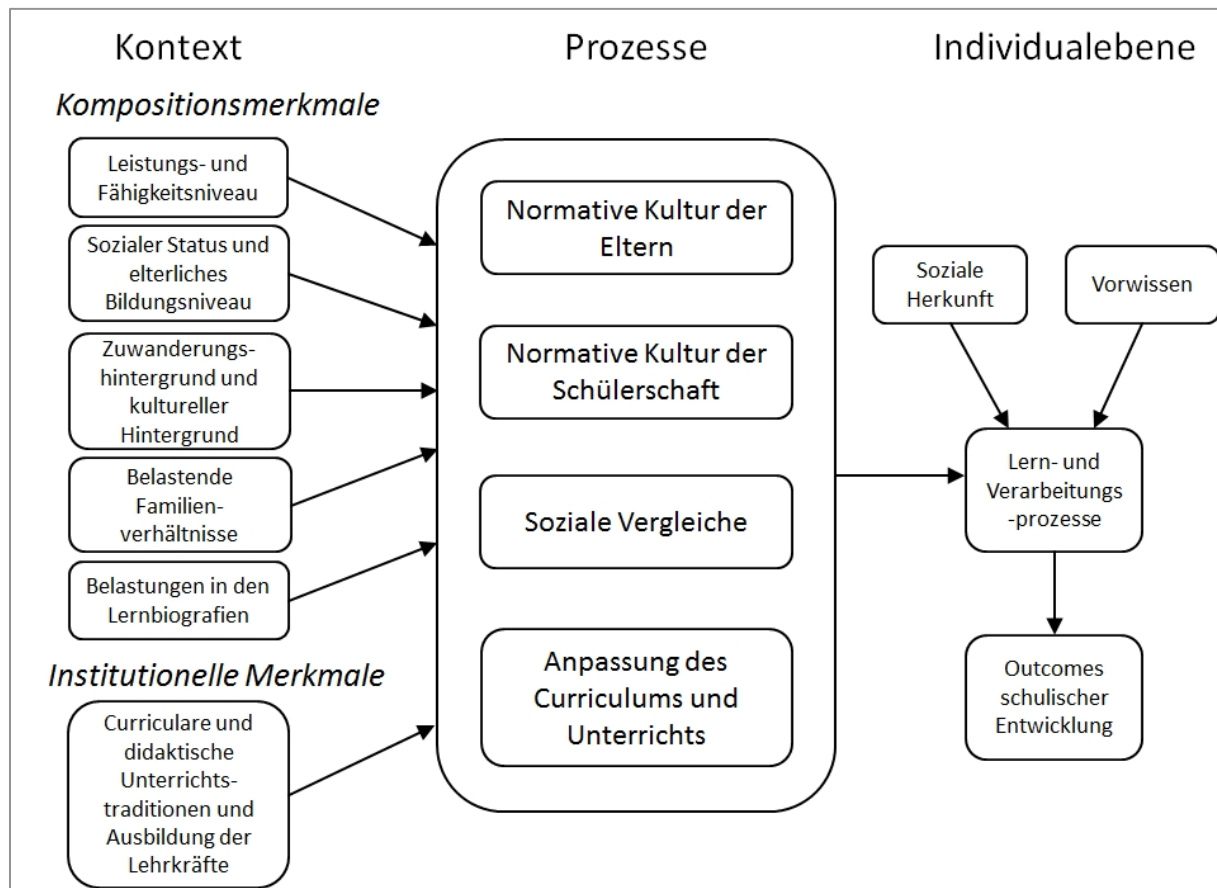


Abbildung 5. Vermittlungsmodell von Effekten der Komposition und Institution auf die individuelle schulische Entwicklung (nach Baumert et al., 2006; S. 126).

Anmerkung: In Baumert et al. (2006, S. 126) werden spezifische Wirkrichtungen von Merkmalen der Komposition auf bestimmte Prozesse angenommen. Zugunsten der Übersichtlichkeit wurden diese spezifischen Pfade nicht abgebildet.

Erstens können Kompositionseffekte über die normative Kultur des Elternhauses vermittelt werden. Dazu zählen elterliche Leistungs- und Verhaltenserwartungen sowie die Unterstützung schulischer Aktivitäten durch die Eltern beziehungsweise Erziehungsberechtigten (Baumert et al., 2006). Die Kooperation der Eltern mit der Schule kann sich durch die Bereitstellung zusätzlicher Anregungen und kultureller Ressourcen (z. B. Museen, Sportvereine etc.) auf die Entwicklung der Schülerinnen und Schüler auswirken. Zusätzlich zu individuellen Einflüssen der elterlichen Ressourcen auf die schulische Entwicklung kann die Zusammensetzung elterlicher Ressourcen in einer Schule in Form von geteilten kulturellen und materiellen Ressourcen die Entwicklung der Schülerschaft beeinflussen.

Ein zweiter Vermittlungsweg ist die normative Kultur der Lerngruppe (z. B. Einstellung zum Lernen, Verhaltensweisen im Unterricht und Akzeptanz von Leistungsnormen, Schulbindung,

s. Abbildung 5). So könnte ein hoher Anteil von Schülerinnen und Schülern mit familiären und lernbiografischen Belastungen zu einem Lernklima beitragen, das sich durch geringe Erfolgserwartungen und einer geringen Wertschätzung von akademischem Fortschritt auszeichnet. Ein solches Lernklima könnte sich negativ auf die individuelle Entwicklung auswirken (Agirdag, Van Houtte & Van Avermaet, 2012).

Drittens werden Vergleichsprozesse zwischen Schülerinnen und Schülern als Vermittlungsweg von Kompositionseffekten betrachtet. Dreeben und Barr (1988) vermuten, dass soziale Bezugsgruppeneffekte die Selbst- und Fremdeinschätzung der Schülerinnen und Schüler, das Tempo und die Schwierigkeit des Unterrichts beeinflussen. Dabei differenzieren Dreeben und Barr (1988) zwischen normativen und komparativen Vergleichsprozessen. Normative Vergleichsprozesse beziehen sich auf die Ausbildung gemeinsamer Verhaltens- und Leistungsnormen, die von einer Lerngruppe geteilt werden und das Unterrichtsklima beeinflussen (Bellin, 2009). Dazu zählen akzeptierte Regeln zum Umgang mit Lehrkräften und geteilte Vorstellungen zum Erreichen von Leistungszielen (Jerusalem, 1997).

Unter komparativen Vergleichsprozessen verstehen Dreeben und Barr (1988) vor allem Leistungsvergleiche zwischen Mitgliedern einer Lerngruppe. Schülerinnen und Schüler vergleichen sich mit ihren Mitschülerinnen und Mitschülern; diese sozialen Vergleichsprozesse beeinflussen schulische Leistungen und motivationale Merkmale (Haeberlin, 1991; Rheinberg & Entstrup, 1977). Soziale Vergleiche mit leistungstärkeren Schülerinnen und Schülern können einerseits die Motivation und Anstrengungsbereitschaft erhöhen und dadurch einen positiven Einfluss auf die Leistungsentwicklung ausüben; andererseits können sie die eigene Leistungseinschätzung verringern, zu Frustration führen und dadurch die Leistungsentwicklung bremsen (s. Kapitel 2.5.3 und Kapitel 2.5.4).

Darüber hinaus können auch soziale Lernprozesse zwischen Schülerinnen und Schülern einer Lerngruppe zu Kompositionseffekten beitragen (z. B. *Peer-Tutoring*; Bandura, 1971; Wilkinson, 2002). Soziale Lernprozesse sind besonders effektiv, wenn Novizen von Experten lernen (Azmitia, 1988). Für leistungsschwächere Schülerinnen und Schüler sollen demnach Interaktionen mit leistungstärkeren Mitschülerinnen und Mitschülern leistungsförderlich sein, da sie Lernstrategien beobachten, imitieren und internalisieren können (Köller, 2004). Außerdem können soziale Lernprozesse durch positive und unterstützende Beziehungen zwischen Schülerinnen und Schülern sowie Interaktionen zwischen Lehrkräften und der Schülerschaft gefördert werden, die ebenfalls durch die Zusammensetzung der Schülerschaft beeinflusst werden (Anderman, 2002; Deci, Vallerand, Pelletier & Ryan, 1991).

Ein vierter Vermittlungsweg für Kompositionseffekte ist die Anpassung des Unterrichts an die wahrgenommenen Bedürfnisse der Lerngruppe durch die Lehrkräfte. Dieser Vermittlungsweg bezieht sich nicht auf den intendierten Lehrplan, sondern auf das tatsächliche Unterrichtsgeschehen (Harris, 2010; Rumberger & Palardy, 2005). So gestalten Lehrkräfte den Unterricht in leistungstärkeren Lerngruppen anspruchsvoller als in leistungsschwächeren Lerngruppen (Köller et al., 2013). Damit verbunden werden in leistungstärkeren Lerngruppen weniger repetitive Unterrichtselemente eingesetzt und die verfügbare Lernzeit wird effektiver genutzt (Baumert, Roeder, Sang & Schmitz, 1986; Bellin, 2009). Weiterhin nehmen Dreeben und Barr (1988) an, dass eine hohe Anzahl an leistungsschwachen Schülerinnen und Schülern in einer Lerngruppe dazu führen, dass die Mitschülerinnen und Mitschüler stärker abgelenkt werden und die Lehrkraft sich nicht gleichermaßen um den Lernfortschritt aller Schülerinnen und Schüler einer Lerngruppe kümmern kann.

2.5.2.3 Institutionelle Effekte

Als institutionelle Effekte werden Einflüsse der Lernumwelt bezeichnet, die auf organisatorische Aspekte, curriculare Unterschiede sowie Merkmale der Lehrkräfte zurückzuführen sind und zusätzlich zu den Effekten der Schülerzusammensetzung wirken (Maaz, Trautwein, Lüdtke, Baumert, 2008). Gamoran (1986) unterscheidet zwei miteinander verbundene Merkmale, die zu institutionellen Effekten in Folge von Leistungsgruppierungen beitragen können: die Unterrichtsgestaltung (*instructional effects*) und die schulische Organisation (*institutional effects*).

Effekte der Unterrichtsgestaltung (*instructional effects*) beziehen sich auf Unterschiede in der Quantität und Qualität des Unterrichts zwischen Schularten beziehungsweise leistungsdifferenzierten Lerngruppen (Gamoran, 1986). Quantitative Merkmale des Unterrichts, wie die Anzahl der Wochenstunden in einem bestimmten Fach und die effektiv zur Verfügung stehende Lernzeit, können Lern- und Vergessensprozesse der Schülerinnen und Schüler beeinflussen (Köller & Baumert, 2012). In leistungstärkeren Lerngruppen werden gleiche Lerninhalte zudem schneller bearbeitet als in leistungsschwächeren Lerngruppen (Gamoran, 1986). Qualitative Merkmale des Unterrichts umfassen die durch das Curriculum bereitgestellten Lerngelegenheiten, die Anpassung der Unterrichtsmethoden an die Lernvoraussetzung der Lerngruppe und Merkmale der Lehrkraft (Creemers & Reezigt, 1996). Ein leistungsförderlicher Unterricht lässt sich zusammenfassend durch ein hohes Anforderungsniveau mit

starkem Fokus auf das Curriculum, der Bereitstellung individueller Hilfe, klare Instruktionen und eine gute Klassenführung charakterisieren (Helmke, 2009). Weiterhin können Unterschiede in der Klassenführung, wie zum Beispiel der Umgang mit Störungen, zu einer differenziellen Leistungsentwicklung der Schülerinnen und Schüler führen (Köller & Baumert, 2012). Auch das fachliche, didaktische und fachdidaktische Wissen der Lehrkraft trägt zu institutionellen Effekten bei (Krauss et al., 2008). Die Ausbildung der Lehrkräfte erfolgt in Deutschland schulartspezifisch und ist insbesondere im Vergleich zwischen Förderschulen und allgemeinen Schulen darauf ausgerichtet, unterschiedliche Vorkenntnisse, Leistungsstärken und Bedürfnissen der Lerngruppe angemessen zu berücksichtigen (Weis-
haupt, 2017).

Effekte der schulischen Organisation (*institutional effects*) beziehen sich auf schulartspezifische Normen, Erwartungen und Statuswahrnehmungen der schulischen Akteure und Eltern (Gamoran, 1986; Meyer, 1977). Merkmale der schulischen Organisation wirken sich insbesondere dann auf die schulische Entwicklung aus, wenn die Zuweisung zu unterschiedlichen Schularten beziehungsweise Lerngruppen mit einem unterschiedlichen Status hinsichtlich der zukünftigen sozialen Rollen ihrer Schülerklientel verbunden ist (Meyer, 1977). Nach Gamoran (1986) wirken sich Merkmale der schulischen Organisation somit indirekt auf die schulische Entwicklung aus, indem sie Erwartungen und Anspruchshaltungen schulischer Akteure beeinflussen. Lehrkräfte in leistungsstärkeren Lerngruppen entwickeln zum Beispiel höhere Leistungserwartungen, die sich wiederum positiv auf die Leistungsentwicklung auswirken (Diamond et al., 2004; Rist, 2000). Gamoran (1986) sowie Pallas, Entwisle, Alexander und Stuka (1994) stellten fest, dass bereits zu Beginn der Schullaufbahn Kinder in leistungsstärkeren Lerngruppen positivere Prognosen bezüglich ihrer akademischen Entwicklung erhalten.

Mit zunehmendem Alter werden die Erwartungen von Eltern und Lehrkräften in die Selbsteinschätzung der Schülerinnen und Schüler integriert und wirken sich auf deren Erfolgserwartungen und schulische Motivation aus (Gamoran, 1986). Hier werden insbesondere die Forschungsansätze von Goffman (1963) zum Labeling beziehungsweise zur Stigmatisierung als Vermittlungsprozesse aufgegriffen. Insbesondere bei Schülerinnen und Schülern in Haupt- und Förderschulen wird angenommen, dass das geringe Ansehen dieser Schulart von ihnen wahrgenommen und internalisiert wird (Knigge, 2017; Pfahl, 2011). Als Folge sollten diese Schülerinnen und Schüler geringere Erfolgserwartungen berichten, was sich im Sinne einer

selbsterfüllenden Prophezeiung (Brophy, 1983; Rosenthal & Jacobson, 1968) auf die Leistungsentwicklung auswirken kann.

Die Bedeutung institutioneller Effekte für die individuelle Entwicklung von Kindern und Jugendlichen wurde in der Bildungsforschung lange Zeit als eher gering eingeschätzt (Köller & Baumert, 2012). Auch aktuelle Meta-Analysen weisen auf eher geringe Einflüsse der schulischen Organisation auf die schulische Entwicklung von Kindern und Jugendlichen hin (Hattie, 2008). Dieses Bild ändert sich jedoch, wenn man zu den institutionellen Effekten nicht nur Merkmale der schulischen Organisation und der Ausstattung zählt, sondern auch systematische Unterschiede zwischen Schularten in der Unterrichtsgestaltung und der Ausbildung der Lehrkräfte betrachtet (Hattie, 2008). So schlussfolgern Köller und Baumert (2001) aus ihrer Studie zu Schülerinnen und Schülern ohne SPF in der Sekundarstufe I, dass Merkmale der Unterrichtsgestaltung und Unterrichtskultur für die Leistungsentwicklung einen bedeutsameren Einfluss haben als Merkmale der Schülerzusammensetzung. Da die Zuweisung der Schülerinnen und Schüler zu unterschiedlichen Schularten aber auch zu differenziellen Schülerzusammensetzungen führt (s. Kapitel 2.5.2.1), ist eine Trennung institutioneller und kompositioneller Effekte in empirischen Studien schwierig (Baumert et al., 2006).

2.5.2.4 Interaktionen zwischen kompositionellen und institutionellen Effekten

Institutionelle und kompositionelle Effekte wirken nicht isoliert voneinander, sondern stehen in Wechselwirkung (Maaz et al., 2008). Die Lernumgebung im Klassenzimmer ist ein dynamisches System (Ennis, 1992): Lehrkräfte erhalten durch ihre Ausbildung sowie institutionell verankerte Lehrpläne und Unterrichtsmethoden das Rüstzeug für die Unterrichtsgestaltung. Anschließend richten sie ihren Unterricht an der wahrgenommenen Leistungsstärke der Schülerinnen und Schüler aus, die wiederum auf die Unterrichtsgestaltung reagieren. Dieses Feedback führt zu Anpassungen durch die Lehrkräfte und erneuten Reaktionen der Lerngruppe (Pallas et al., 1994). Zudem besteht in leistungsdifferenzierenden, gegliederten Bildungssystemen ein starker Zusammenhang zwischen institutionellen Merkmalen der Unterrichts- und Schulorganisation und kompositionellen Merkmalen der Schülerzusammensetzung (Baumert et al., 2006). Das liegt vor allem daran, dass Kinder und Jugendliche auf Basis ihrer vorherigen Leistungen zu Schularten mit unterschiedlichen Curricula und Unterrichtsorganisationen zugewiesen werden.

Darüber hinaus kann sich die Schülerzusammensetzung in Abhängigkeit von der Schulart unterschiedlich auf die individuelle schulische Entwicklung auswirken. So fanden Baumert et al. (2006) innerhalb des Gymnasiums keine bedeutsamen Zusammenhänge zwischen dem mittleren Fähigkeitsniveau einer Lerngruppe und der individuellen Lesekompetenz. Baumert et al. (2006) führen dieses Ergebnis auch auf die stabil hohen Leistungserwartungen in dieser Schulart zurück, die sich sowohl in den Lerninhalten als auch in der Erwartungshaltung der Lehrkräfte widerspiegeln. Im Gegensatz dazu zeigte sich in Hauptschulen der stärkste Zusammenhang zwischen der leistungsbezogenen Schülerzusammensetzung und der individuellen Leistung im Leseverständnis (Baumert et al., 2006). Hier scheint ein erfolgreicher Unterricht am stärksten von der fähigkeitsbezogenen Schülerzusammensetzung abzuhängen.

Interaktionen zwischen institutionellen Merkmalen der Schulart und Merkmalen der Schülerzusammensetzung können auf mindestens zwei Ursachen zurückgeführt werden. Erstens können Merkmale der Schülerzusammensetzung zwischen den Schularten unterschiedlich stark streuen. So variiert zum Beispiel der Anteil an Schülerinnen und Schülern mit belastenden Lernbiografien innerhalb von Gymnasien weniger stark als innerhalb von Hauptschulen (Baumert et al., 2006). Dadurch ist es zum einen schwieriger, Effekte dieses Kompositionsmerkmals innerhalb von Gymnasien festzustellen. Zum anderen könnte auch der absolute Anteil an leistungsschwachen Schülerinnen und Schülern so gering sein, dass er für die Unterrichtsgestaltung praktisch nicht relevant ist (Baumert et al., 2006). Im Gegensatz dazu ist der Anteil leistungsschwacher Schülerinnen und Schüler in Hauptschulen deutlich höher und könnte die Unterrichtsgestaltung in dieser Schulart deshalb stärker beeinflussen. Zweitens könnten Lehrkräfte je nach Schulart unterschiedlich flexibel auf Veränderungen der Schülerzusammensetzung reagieren (Baumert et al., 2006). Beispielsweise könnte eine gesunkene Nachfrage am Besuch einer Hauptschule (bei steigender Nachfrage eines Gymnasialbesuchs) dazu beitragen, dass vor allem Schülerinnen und Schüler mit belastenden Lernbiografien an Hauptschulen unterrichtet werden (Bos et al., 2010; Dumont et al., 2013). Falls die Aus- und Fortbildung der Hauptschullehrkräfte an die veränderte Schülerschaft nicht so effektiv angepasst werden kann wie die Ausbildung der Gymnasiallehrkräfte, würde das ebenfalls Wechselwirkungen zwischen institutionellen und kompositionellen Merkmalen zur Folge haben (Baumert et al., 2006).

2.5.3 Schulische Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen

Ausgehend davon, dass Schularten differenzielle Lern- und Entwicklungsmöglichkeiten bereitstellen können, wird in diesem Kapitel diskutiert, inwieweit sich die schulische Entwicklung von Kindern und Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen unterscheidet. Dabei liegt der Fokus auf den schulischen Kompetenzen von Schülerinnen und Schülern mit SPF. Schulische Kompetenzen sind das Ergebnis individueller Lernprozesse, die domänenspezifisch erworben und von Lerngelegenheiten in der Schule beeinflusst werden (Köller, 2010). Zunächst werden Argumente vorgestellt, warum Kinder und Jugendliche mit SPF in Förderschulen oder allgemeinen Schulen höhere schulische Kompetenzen erzielen könnten. Anschließend wird der Forschungsstand zum Einfluss der Beschulungsart auf die schulischen Leistungen von Kindern und Jugendlichen mit SPF zusammengefasst.

2.5.3.1 Argumente für höhere schulische Kompetenzen in Förderschulen

Ein zentrales Argument für die positive Auswirkung von Förderschulen auf die Leistungsentwicklung ist, dass Förderschulen einen Schonraum für Kinder und Jugendliche mit SPF darstellen (Schumann, 2007). Dieses Argument bezieht sich zum einen auf die Zusammensetzung der Lerngruppen in Förderschulen und zum anderen auf institutionelle Merkmale der Förderschulen.

Zum einen lernen Kinder und Jugendliche mit SPF in Förderschulen gemeinsam mit Schülerinnen und Schülern mit ähnlichen Lernproblemen. Dadurch seien sie von zu hohen Leistungsansprüchen, Leistungsdruck, schulischem Versagen und auch sozialer Ausgrenzung geschützt (Tent, Witt, Bürger & Zschoche-Lieberum, 1991). Diese geschützte Lernumgebung in Förderschulen könnte sich positiv auf die Leistungsentwicklung von Kindern und Jugendlichen mit SPF auswirken (Haeberlin, 1991). Weiterhin werden in Förderschulen kleinere Lerngruppen gebildet als in allgemeinen Schulen (Schor, 2003; Wocken, 2000, 2005). In kleineren Lerngruppen könnten Lehrkräfte stärker auf die individuellen Bedürfnisse der Schülerinnen und Schüler eingehen (Pianta, 1999). Außerdem wird angenommen, dass es Lehrkräften in kleineren, leistungshomogenen Lerngruppen leichter falle, ihren Unterricht auf die Lerngruppe auszurichten (Helmke & Weinert, 1997). In Förderschulen lasse sich demnach das Unterrichtstempo und das Anforderungsniveau stärker auf die homogene Schülergruppe

mit SPF anpassen, was zu einer Verringerung von Über- und Unterforderungen führen sollte (Ahrbeck, 2014). Außerdem könnten Lehrkräfte in kleineren Lerngruppen in Förderschulen effektiver auf Unterrichtsstörungen reagieren, wodurch die effektive Lernzeit erhöht wird (Vannest & Hagan-Burke, 2010; Vannest & Parker, 2010).

Zum anderen bezieht sich die Annahme einer geschützten Lernumgebung in Förderschulen auf institutionelle Merkmale. Es wird angenommen, dass Schülerinnen und Schüler mit SPF qualifizierte Lehrkräfte und eine spezielle Unterrichtsgestaltung benötigen, um gute Lernfortschritte erzielen zu können (Ahrbeck, 2014). Lehrkräfte in Förderschulen erhalten eine sonderpädagogische Ausbildung, was sie in die Lage versetzen soll, die Leistungsentwicklung von Schülerinnen und Schülern mit SPF besser zu fördern als Lehrkräfte in allgemeinen Schulen (Ahrbeck, 2014, 2016). Außerdem seien die Lehrpläne und Unterrichtsmaterialien in Förderschulen stärker auf die Lernvoraussetzungen von Kindern und Jugendlichen mit SPF ausgerichtet (Vaughn et al., 1996). Die Lehrpläne in Förderschulen hätten unter anderem einen geringeren Umfang an Lerninhalten (Wocken, 2005). Dadurch stehe Lehrkräften in Förderschulen mehr Zeit zum Erklären und zur Wiederholung der Lerninhalte zur Verfügung (Schumann, 2007).

Außerdem würden Förderschulen über mehr Ressourcen zur optimalen Förderung von Schülerinnen und Schülern mit SPF als allgemeine Schulen verfügen (Ahrbeck, 2014). Die schulische Ausstattung in Förderschulen sei stärker auf die Lernbedürfnisse von Schülerinnen und Schülern mit SPF ausgerichtet und es stehen zusätzliche sonderpädagogische Fachkräfte für individuelle Fördermaßnahmen zur Verfügung (Ahrbeck, 2014). Im Gegensatz dazu hätten viele allgemeine Schulen ihre schulische Organisation und Unterrichtsgestaltung (noch) nicht auf die Lernbedürfnisse von Schülerinnen und Schülern mit SPF angepasst (Grosche & Vock, im Druck). Die Abkehr von Förderschulen könnte deshalb insgesamt zu einer Verringerung der Unterrichtsqualität führen, die zu Lasten der Schülerinnen und Schüler mit SPF gehen würde (Ahrbeck, 2014).

2.5.3.2 Argumente für höhere schulische Kompetenzen in allgemeinen Schulen

Argumente für eine positive Auswirkung der allgemeinen Schulen auf die Leistungsentwicklung von Kindern und Jugendlichen mit SPF beziehen sich ebenfalls zum einen auf Merkmale der Schülerzusammensetzung und zum anderen auf institutionelle Merkmale der Unterrichtsorganisation.

Zum einen könnten Schülerinnen und Schüler mit SPF von leistungsstärkeren Mitschülerinnen und Mitschülern in allgemeinen Schulen profitieren (Gersten et al., 2009; Justice, Logan, Lin & Kaderavek, 2014). Es wird angenommen, dass Kinder und Jugendliche mit SPF in allgemeinen Schulen die Lösungsstrategien kompetenterer Kinder und Jugendlicher übernehmen und durch soziale Lernprozesse in ihrer Leistungsentwicklung profitieren (Gersten et al., 2009; Scruggs, Mastropieri & Marshak, 2012). Außerdem schaffe die Anwesenheit positiver Lernvorbilder in allgemeinen Schulen zusätzliche Lernanreize, die sich positiv auf die Leistungsentwicklung auswirken können (Frey, Dauenheimer, Parge & Haisch, 1993; Wocken, 2000). Wocken (2005) verweist darauf, dass Mitschülerinnen und Mitschüler ohne SPF, die bereits eine Zone der nächsten Entwicklung (Wygotski, 1978) erreicht haben, wertvolle Lernimpulse für Kinder und Jugendliche mit SPF geben. Im Gegensatz dazu wird in Förderschulen ein hoher Anteil an Schülerinnen und Schülern mit belastenden Lernbiografien und Misserfolgserfahrungen unterrichtet (s. Kapitel 2.5.2.1). Dadurch könnte in Förderschulen ein ungünstiges Lernklima bestehen, das durch eine geringere kognitive Anregung durch Mitschülerinnen und Mitschüler gekennzeichnet sei (Wocken, 2005).

Zum anderen könnten institutionelle Merkmale zu höheren schulischen Kompetenzen von Schülerinnen und Schülern mit SPF in allgemeinen Schulen beitragen. Die Lehrpläne in allgemeinen Schulen seien umfangreicher und die Unterrichtsgestaltung sei anspruchsvoller als in Förderschulen (Markussen, 2004; Wocken, 2005). Die Unterrichtsgestaltung in allgemeinen Schulen sei stärker auf den akademischen Fortschritt der Schülerinnen und Schüler ausgerichtet (Hocutt, 1996; Schumann, 2007; Wocken, 2000, 2005). So kämen im Unterricht in allgemeinen Schulen anspruchsvollere Aufgaben zum Einsatz, die auch die Kompetenzentwicklung von Kindern mit SPF stärker fördern könnten (Tillmann & Wischer, 2006). Leistungsschwächere Schülerinnen und Schüler würden in ihren schulischen Leistungen von einem Unterrichtsniveau profitieren, welches etwas über ihrem Fähigkeitsniveau liegt (Dreeben & Barr, 1988). Außerdem wird angenommen, dass Lehrkräfte in allgemeinen

Schulen vielfältigere Unterrichtsmethoden einsetzen, um den verschiedenen Leistungsstärken der Schülerinnen und Schüler gerecht zu werden.

Im Gegensatz dazu seien die Lehrpläne in Förderschulen hinsichtlich Umfang und Komplexität deutlich reduziert (Wocken, 2005). Der Unterricht würde stärker auf die Persönlichkeitsentwicklung als auf die kognitive Entwicklung der Schülerinnen und Schüler ausgerichtet (Wocken, 2005). Lerninhalte würden in Förderschulen im Vergleich zu allgemeinen Schulen in kleinere Einheiten herunter gebrochen (Werning & Lütje-Klose, 2006). Dadurch würden Schülerinnen und Schüler mit SPF jedoch weniger lernen, größere Sinnzusammenhänge zu bilden und logisch-deduktive Denkprozesse weiterzuentwickeln (Wocken, 2005). Außerdem basiere der Unterricht in Förderschulen auf vielen wiederholenden Elementen und einer geringeren methodischen Vielfalt als in allgemeinen Schulen (Ellger-Rüttgardt, 2008; Werning & Lütje-Klose, 2006; Wocken, 2005). Deshalb könnten Förderschulen im Vergleich zu allgemeinen Schulen anregungsärmere Lernumgebungen darstellen, in denen Kinder und Jugendliche mit SPF weniger herausgefordert werden (Wocken, 2005).

Als weiteres Argument für die Beschulung von Kindern mit SPF in allgemeinen Schulen wird angeführt, dass die Schaffung spezieller Klassen und Schulen für leistungsschwache Kinder mit einer Stigmatisierung und weitreichenden psychosozialen Konsequenzen für die betroffenen Schülerinnen und Schüler einhergehe (Dunn, 1968; Pfahl & Powell, 2011), was sich negativ auf die schulischen Leistungen dieser Schülergruppe auswirken kann (Banerji & Dailey, 1995; Pfahl, 2012). Im Gegensatz dazu könnten Kinder und Jugendlichen mit SPF in allgemeinen Schulen von der höheren Erwartungshaltung der Lehrkräfte profitieren (Boaler & William, 2001; Gamoran, 1986; Ready & Wright, 2011).

2.5.3.3 Empirische Befunde zum Einfluss des Förderortes auf die schulischen Kompetenzen von Kindern mit SPF

In der Mehrzahl der Untersuchungen zum Einfluss des Förderortes auf die Schulleistungen von Kindern mit SPF zeigten sich neutrale bis positive Effekte zugunsten einer Beschulung dieser Schülerinnen und Schüler in allgemeinen Schulen. Ruijs und Peetsma (2009) kamen nach Sichtung und Zusammenstellung internationaler Studien zur Wirkung integrativer und inklusiver Schulformen zu dem Schluss, dass der Unterricht von Kindern mit SPF in allgemeinen Schulen keine negativen Auswirkungen auf deren schulische Leistungen im sprachli-

chen und mathematischen Bereich habe. Vielmehr könne im Durchschnitt mit geringen positiven Effekten zugunsten einer Beschulung von Kindern mit SPF in allgemeinen Schulen gerechnet werden. Auch ältere Metaanalysen kamen zu dem Schluss, dass die schulischen Leistungen von Kindern mit SPF im sprachlichen und mathematischen Bereich in allgemeinen Schulen höher ausgeprägt waren als in Förderschulen. Die Höhe dieser Effekte variierte dabei zwischen geringen (Baker, Wang & Walberg, 1994; Carlberg & Kavale, 1980) bis mittelstarken Effekten zugunsten der Beschulung in allgemeinen Schulen (Wang & Baker, 1985). Vereinzelt Studien weisen allerdings auch auf negative Effekte inklusiver Beschulung für Schülerinnen und Schüler mit Lernbeeinträchtigungen hin (Harrington, 1997; Rogers & Thiery, 2003).

Der empirische Forschungsstand in Deutschland erlaubt noch kein abschließendes Urteil über die Auswirkung der Beschulungsart auf die schulischen Kompetenzen von Schülerinnen und Schülern mit SPF, da bisher noch nicht ausreichend methodisch anspruchsvolle Studien mit umfangreichen Stichproben durchgeführt wurden (Grosche & Vock, im Druck). Übereinstimmend mit internationalen Studienergebnissen weist jedoch die Mehrzahl der bisherigen Studien auf Lernvorteile von Kindern und Jugendlichen mit SPF in allgemeinen Schulen hin (Werning & Reiser, 2008).

In einer frühen Studie mit hessischen Schülerinnen und Schülern mit dem Förderschwerpunkt *Lernen* konnten Tent, Witt, Zschoche-Lieberum und Bürger (1991) die Schonraumwirkung der Förderschule bestätigen. Dazu zählt eine geringere Prüfungsangst, ein besseres Arbeitsverhalten und ein stärker ausgeprägtes Selbstwertgefühl der Schülerinnen und Schüler in Förderschulen im Vergleich zu allgemeinen Schulen. Trotzdem erzielten Schülerinnen und Schüler mit SPF in Förderschulen unter Kontrolle der kognitiven Leistungsfähigkeit geringere schulische Leistungen als Schülerinnen und Schüler mit SPF in allgemeinen Schulen. Auch Haeberlin (1991) fand höhere schulische Leistungen für Schülerinnen und Schüler mit SPF, die in allgemeinen Schulen unterrichtet wurden.

Die Forschergruppe um Wocken (2000, 2005; Wocken & Gröhlich, 2009) kam im Rahmen mehrerer querschnittlich angelegter Studien zu dem Schluss, dass Schülerinnen und Schüler mit SPF in Förderschulen in ihren schulischen Kompetenzen zunehmend hinter der Schülergruppe ohne SPF zurückfallen. Wocken (2000, 2005) untersuchte die Leistung, Intelligenz und soziale Lage von Schülerinnen und Schülern mit dem Förderschwerpunkt *Lernen*. Dabei wurden in seiner Studie zur Lernausgangslage an Förderschulen (LAUF) nahezu alle Förder-

schulen mit dem Schwerpunkt *Lernen* in Hamburg in die Stichprobe aufgenommen und mit einer zusätzlichen Stichprobe an Förderschulen in Brandenburg und Niedersachsen verglichen. Allerdings konnten keine Schülerinnen und Schüler mit SPF in allgemeinen Schulen als Vergleichsgruppe einbezogen werden. Dieser Limitation wurde durch zwei Kontrollvariablen begegnet: Erstens wurde die Dauer der Beschulung in Förderschulen (in Schuljahren) als zusätzlicher Prädiktor der schulischen Leistungen aufgenommen und zweitens wurden Schülerinnen und Schüler in Hauptschulen als Vergleichsgruppe gewählt. Mit zunehmender Beschulungsdauer in der Förderschule fielen Schülerinnen und Schüler in Förderschulen in ihren Leistungsständen im Rechtschreiben und in ihren kognitiven Fähigkeiten im Vergleich zu Schülerinnen und Schülern in Hauptschulen immer weiter zurück (Wocken, 2005). Wocken (2005) schlussfolgert daraus, dass Förderschulen die Leistungsrückstände von Schülerinnen und Schülern mit SPF nicht angemessen kompensieren können.

Im Gegensatz dazu kamen Lehmann und Hoffmann (2009) in einer querschnittlich angelegten Untersuchung von Berliner Jugendlichen mit dem Förderschwerpunkt *Lernen* in den Klassen 7 bis 10 zu dem Schluss, dass Jugendliche in Förderschulen über verschiedene Schulfächer und berufsrelevante Kompetenzen hinweg ähnlich abschnitten wie Jugendliche mit SPF in allgemeinen Schulen. So waren die mathematischen und sprachlichen Leistungen von Jugendlichen mit SPF in allgemeinen Schulen in der 7. und 8. Jahrgangsstufe höher als die von Jugendlichen in Förderschulen; in der 9. und 10. Klasse zeigten sich jedoch Leistungsvorteile zugunsten der Förderschülerinnen und -schüler. Allerdings lässt sich als eine zentrale Einschränkung anmerken, dass die Umkehrung der Befunde in der 9. und 10. Jahrgangsstufe auch darauf zurückgeführt werden könnten, dass leistungsstarke Jugendliche mit SPF in allgemeinen Schulen ihren Förderstatus zum Ende der Schulkarriere verlieren, um einen regulären Schulabschluss erwerben zu können (Preuss-Lausitz, 2009). Dadurch dürften in der Studie in allgemeinen Schulen eher leistungsschwächere Schülerinnen und Schüler in den Vergleich zwischen den Schularten eingegangen sein, was die Ergebnisse verzerrt haben könnte (Preuss-Lausitz, 2009).

In der Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements (BiLieF, Wild et al., 2015) wurde die kognitive, motivationale und psychosoziale Entwicklung von Kindern mit dem Förderschwerpunkt *Lernen* untersucht. Die Studie begleitete Kinder mit SPF in Förderschulen und allgemeinen Schulen in Nordrhein-Westfalen über die Grundschulzeit. Stranghöner et al. (2017) verglichen auf Grundlage von Daten der BiLieF-Studie Kinder mit SPF in Förderschulen und allgemeinen Schulen hinsichtlich ihrer

Leistungsentwicklung im Lesen und im Rechtschreiben. Dabei wurden Daten von insgesamt 461 Kindern über drei Messzeitpunkte mit latenten Wachstumskurvenanalysen ausgewertet. Unter Berücksichtigung der kognitiven Fähigkeiten, der sozialen Herkunft und des Geschlechts der Schülerinnen und Schüler zeigten sich deutlich höhere Ausgangsleistungen im Lesen und Rechtschreiben bei Kindern mit SPF in allgemeinen Schulen. In beiden Schularten erzielten Kinder mit SPF bedeutsame Leistungszuwächse im Lesen und Rechtschreiben. Während diese Zuwächse im Bereich Lesen für Kinder mit SPF in Förderschulen und allgemeinen Schulen vergleichbar ausfielen, erzielten Kinder mit SPF in Förderschulen höhere Leistungszugewinne im Rechtschreiben als Kinder mit SPF in allgemeinen Schulen. Berücksichtigt man die Ausgangsleistungen zu Beginn der dritten Klasse, dann erzielten Kinder mit dem Förderschwerpunkt *Lernen* in allgemeinen Schulen höhere schulische Leistungen als Kinder in Förderschulen. Diese Unterschiede bestehen im Verlauf der Grundschulzeit bis zum Ende der vierten Klasse fort, auch wenn Kinder mit SPF in Förderschulen im Bereich des Rechtschreibens aufholten (Stranghöner et al., 2017).

Der deutschsprachige Forschungsstand zum Einfluss der Beschulungsart auf die schulischen Kompetenzen von Schülerinnen und Schüler mit SPF erlaubt noch kein abschließendes Urteil (Grosche & Vock, im Druck). Wesentliche Limitationen der vorliegenden Studien bestehen darin, dass nur selten standardisierte Leistungstests eingesetzt wurden und adäquate Kontrollgruppen zur Verfügung standen. Weiterhin wurden nur wenige Hintergrundmerkmale der Schülerinnen und Schüler in den Analysen berücksichtigt, so dass die berichteten Effekte der Schulart auf die schulischen Leistungen zum Teil auf bereits bestehende Unterschiede zwischen den Schülergruppen zurückgehen könnten. Bisherige Studien fokussieren zudem stark auf die Gruppe der Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen*, nach Förderschwerpunkt differenzierte Befunde liegen nicht vor.

2.5.4 Schulische Motivation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen

Neben der Förderung der kognitiven Entwicklung sind die motivationale Entwicklung und Persönlichkeitsentwicklung weitere zentrale Bildungs- beziehungsweise Erziehungsziele der Schule (Baumert et al., 2006; Hofer, 2014). Diese Merkmale werden von der Lernumgebung in der besuchten Schule beziehungsweise Schulart beeinflusst (Urda & Schoenfelder, 2006; Wang & Eccles, 2013). Wiederum können sich einerseits schulartspezifische Schülerzusammensetzungen auf die individuelle schulische Motivation auswirken (Allodi, 2000; Dijkstra, Kuyper, van der Werf, Buunk & van der Zee, 2008; Marsh, 2005; Trautwein et al., 2006). Andererseits stehen institutionelle Merkmale, wie zum Beispiel Leistungserwartungen, Formen der Leistungsrückmeldung und die Unterrichtsgestaltung, in Zusammenhang mit der motivationalen Entwicklung der Schülerinnen und Schüler (Baumert et al., 2006; Lüdtke, Köller, Marsh & Trautwein, 2005; Ryan & Deci, 2000).

Im Folgenden werden mit dem akademischen Selbstkonzept und der Lernfreude zwei zentrale Merkmale der schulischen Motivation beschrieben. Anschließend wird auf zwei Mechanismen eingegangen, die zu Unterschieden in der schulischen Motivation zwischen Kindern und Jugendlichen verschiedener Schularten beitragen können. Darauf aufbauend folgt eine Gegenüberstellung von Argumenten, die für eine höhere schulische Motivation bei Schülerinnen und Schülern mit SPF in Förderschulen oder für eine höhere schulische Motivation bei Schülerinnen und Schülern mit SPF in allgemeinen Schulen sprechen. Das Kapitel schließt mit einer Zusammenfassung des Forschungsstandes zum Zusammenhang zwischen Beschulungsart und Merkmalen der schulischen Motivation von Kindern und Jugendlichen mit SPF.

2.5.4.1 Merkmale der schulischen Motivation: Akademisches Selbstkonzept und Lernfreude

Die schulische Motivation ist insbesondere bei Schülerinnen und Schülern mit SPF ein wichtiger Aspekt ihrer schulischen Entwicklung (Elbaum, 2002; Grolnick & Ryan, 1990; Möller, Streblow & Pohlmann, 2009). Schülerinnen und Schüler mit SPF erleben im schulischen Alltag häufiger Misserfolge als Schülerinnen und Schüler ohne SPF, was sich negativ auf ihre schulische Motivation auswirken kann (Allodi, 2000; Grolnick & Ryan, 1990). So kann ein Teufelskreislauf entstehen, bei dem Misserfolgserfahrungen zu einem Absenken der Motivation führen, was wiederum zu niedrigeren schulischen Leistungen führen kann

(Peetsma et al., 2001). Schülerinnen und Schüler mit SPF, die wiederholte schulische Misserfolge erleben, entwickeln geringere Erfolgserwartungen und nehmen ihre Fähigkeiten eher als unveränderbar wahr (Deci & Chandler, 1986). Dadurch könnten sie sich bei zukünftigen Lernaufgaben oder Tests weniger anstrengen und das Erreichen von Lernzielen als weniger wichtig erachten. Das kann sich wiederum negativ auf zukünftige schulische Leistungen auswirken (Deci & Chandler, 1986). Deci und Chandler (1986) sehen darin eine wesentliche Ursache für die Persistenz von Lern- und Leistungsstörungen bei Kindern und Jugendlichen mit SPF.

Die schulische Motivation umfasst viele Merkmale, die sich zum Teil nur schwer von emotionalen Merkmalen abgrenzen lassen (Helmke & Weinert, 1997). Motivationstheorien systematisieren verschiedene Merkmale der schulischen Motivation und setzen sie in Beziehung zu Lernprozessen. Dazu gehört auch das Erwartungs-Wert-Modell von Eccles (Eccles et al. 1983; Wigfield & Eccles, 2000). Wigfield und Eccles (2000) gehen davon aus, dass schulische Leistungen und schulische Entscheidungen durch Erfolgserwartungen und Werturteile der Schülerinnen und Schüler beeinflusst werden. Zu den Erfolgserwartungen gehören Einschätzungen von Kindern und Jugendlichen darüber, wie sie bei einer bestimmten Aufgabe oder in einem bestimmten Fach abschneiden werden (Wigfield & Eccles, 2000; Nagengast et al., 2011). Werturteile umfassen vier verschiedene Bereiche: (1) die intrinsische Lernfreude und das subjektive Interesse an einem Unterrichtsfach, (2) die Bedeutsamkeit, die einer guten Leistung in einem Unterrichtsfach beziehungsweise in einem Test beigemessen wird, (3) der wahrgenommene Nutzen eines Unterrichtsfachs für das Erreichen von zukünftigen Zielen und (4) die wahrgenommenen Kosten, die mit einer schulischen Aktivität verbunden sind. Gemeinsam beeinflussen Erwartungen und Werturteile Anstrengungsbereitschaft, Ausdauer und schulisches Engagement von Kindern und Jugendlichen (Eccles & Wigfield, 2002; Meece, Anderman & Anderman, 2006). Dadurch wirken sich Erwartungen und Werturteile auf die schulischen Leistungen, aber auch auf die Wahl von Leistungskursen und auf Berufswünsche aus (Eccles et al., 1983; Nagengast et al., 2011; Wang & Eccles, 2013).

Das akademische Selbstkonzept ist ein zentraler Bestandteil der Erwartungskomponente und bezieht sich auf Selbsteinschätzungen der Schülerinnen und Schüler und Wahrnehmungen der eigenen schulischen Leistungen in einem bestimmten Fach (Marsh, Kong & Hau, 2000; Trautwein et al., 2006). Es wird durch die Selbsteinschätzung bisheriger schulischer Leistungen gebildet und kann sich dadurch von Fremdbeurteilungen oder Testergebnissen unterscheiden (Shavelson, Hubner & Stanton, 1976). Zwar sind Leistungsrückmeldungen (z. B. in

Form von Schulnoten) Grundlage der Selbsteinschätzung von Kindern und Jugendlichen, jedoch wirken auch soziale Vergleichsprozesse (Dijkstra et al., 2008; Marsh, 2005) sowie Überzeugungen und Stereotype der Eltern und Lehrkräfte (Diamond et al., 2004; Rist, 2000) auf die Ausbildung des akademischen Selbstkonzeptes ein.

Ein zentraler Bestandteil der intrinsischen Wertkomponente ist die Lernfreude der Schülerinnen und Schüler (Wigfield & Eccles, 2000). Darunter wird das subjektive Interesse an neuen Lerninhalten und der Spaß am Unterricht verstanden (Jerusalem et al., 2009). Während das akademische Selbstkonzept überwiegend durch kognitive Einschätzungen gebildet wird, bezieht sich die Lernfreude von Kindern und Jugendlichen stärker auf affektive Prozesse (Pekrun, 1992).

Akademisches Selbstkonzept und Lernfreude beeinflussen sich wechselseitig; entsprechend berichten Schülerinnen und Schüler mit niedrigen akademischen Selbstkonzepten tendenziell auch eine geringere Lernfreude im Unterricht (Wigfield & Eccles, 2000). Beide Merkmale der schulischen Motivation beeinflussen die Leistungsentwicklung von Kindern und Jugendlichen (Valentine, DuBois & Cooper, 2004; Pekrun, Goetz, Titz & Perry, 2002). Schülerinnen und Schüler, die ihre Fähigkeiten positiv einschätzen und Spaß im Unterricht empfinden, initiieren eher leistungsrelevante Tätigkeiten und suchen herausfordernde Lernsituationen auf (Eccles & Wigfield, 2002; Meece et al., 2006). Außerdem begünstigen hohe Ausprägungen des akademischen Selbstkonzepts und der Lernfreude die Fortführung von Lernaktivitäten bei Schwierigkeiten und bei Misserfolgen (Helmke & Weinert, 1997; Pekrun et al., 2002). Dabei bestehen wechselseitige Einflüsse zwischen schulischen Leistungen, akademischem Selbstkonzept und Lernfreude (Marsh et al., 2008; Valentine et al., 2004; Wigfield & Eccles, 2000).

2.5.4.2 Einflüsse der Lernumgebung auf die schulische Motivation

Merkmale der schulischen Motivation von Kindern und Jugendlichen werden durch die Lernumgebung beeinflusst (Urda & Schoenfelder, 2006; Wang & Eccles, 2013). Im Folgenden werden zwei Einflussfaktoren beschrieben, die zu Schulartunterschieden in der schulischen Motivation beitragen: soziale Vergleiche und Unterschiede in der sozialen Integration der Schülerinnen und Schüler.

a) Soziale Vergleiche und der Big-Fish-Little-Pond-Effekt

Die Theorie sozialer Vergleichsprozesse (*Social Comparison Theory*; Festinger, 1954) geht davon aus, dass Leistungsvergleiche zwischen Schülerinnen und Schülern einer Lerngruppe eine wichtige Quelle für die Ausbildung akademischer Selbstkonzepte und verwandter motivationaler Merkmale sind. Demnach vergleichen Schülerinnen und Schüler ihre schulischen Leistungen bei schulischen Aufgaben und ihre Beiträge im Unterricht mit denen ihrer Mitschülerinnen und Mitschüler, um ein akkurates Selbstbild zu erhalten (Marsh et al., 2008). Vergleichsprozesse können darüber hinaus auch initiiert werden, um ein positives Selbstbild aufrechtzuerhalten oder um Leistungsanstrengungen zu erhöhen (Dijkstra et al., 2008).

Marsh (1987) greift die zentrale Bedeutung sozialer Vergleichsprozesse für das akademische Selbstkonzept in seinem Modell der internen und externen Referenzrahmen auf. Zum einen vergleichen Schülerinnen und Schüler ihre aktuellen Leistungen mit vorangegangenen Leistungen (temporale Vergleiche; Marsh, 1986) beziehungsweise in verschiedenen Fächern (dimensionale Vergleiche; Möller & Marsh, 2013). Beide Vergleichsprozesse beinhalten einen internen Referenzrahmen (Marsh, 1987). Zum anderen vergleichen sich Kinder und Jugendliche mit ihren Mitschülerinnen und Mitschülern, einem externen Referenzrahmen. Marsh et al. (2008) nehmen dabei an, dass Schülerinnen und Schüler eine mentale Repräsentation der mittleren Leistungsstärke ihrer Lerngruppe als Vergleichsgrundlage nutzen. Der Einfluss externer Referenzrahmen auf Merkmale der schulischen Motivation wird auch als Fischteicheffekt (*Big-Fish-Little-Pond-Effect*; Marsh, 1987) bezeichnet. Darunter wird verstanden, dass Schülerinnen und Schüler mit vergleichbaren individuellen Leistungen in leistungstärkeren Lerngruppen („große Teiche“) ein niedrigeres Selbstkonzept als in leistungsschwächeren Lerngruppen („kleine Teiche“) berichten (Abbildung 6).

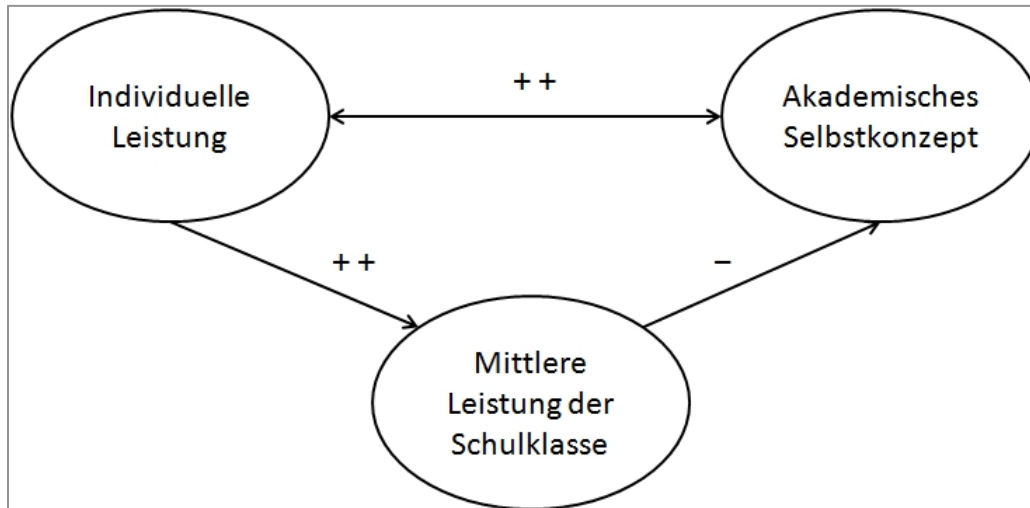


Abbildung 6. Der Big-Fish-Little-Pond-Effekt (Marsh, 1987).

Der negative Einfluss der mittleren Leistungsstärke einer Lerngruppe auf das akademische Selbstkonzept lässt sich auf zwei gegenläufige Effekte zurückführen: Assimilationseffekte und Kontrasteffekte (Corcoran, Crusius & Mussweiler, 2011; Marsh et al., 2008). Auf der einen Seite beziehen sich *Assimilationseffekte* auf die Tendenz von Schülerinnen und Schülern, Selbsteinschätzungen in Übereinstimmung mit der wahrgenommenen Leistungsstärke ihrer Lerngruppe zu bilden (Corcoran et al., 2011; Dijkstra et al., 2008). Demnach solle der Besuch einer prestigeträchtigen Schule oder einer Klasse für leistungsstarke Schülerinnen und Schüler zu positiveren Selbsteinschätzungen führen. Diese Form von Assimilation wird auch als *Basking-in-Reflected-Glory-Effekt* bezeichnet (Cialdini et al., 1976). Im Gegensatz dazu solle der Besuch einer Schule oder Schulart, der eine geringe Leistungsstärke zugeschrieben wird, zu negativeren Selbsteinschätzungen der eigenen Leistungsfähigkeit führen. Diese Form von Assimilation wird als *Labeling-Effekt* oder *Stigmatisierungseffekt* bezeichnet (Goffman, 1963, 1986; Rist, 2000). *Kontrasteffekte* sind das Ergebnis sozialer Vergleichsprozesse, die Unterschiede zwischen einem Kind oder Jugendlichen und der Lerngruppe betonen (Corcoran et al., 2011). Kontrasteffekte treten auf, wenn Schülerinnen und Schüler mit vergleichbaren individuellen Leistungen in leistungsschwächeren Lerngruppen höhere akademische Selbstkonzepte ausbilden als in leistungsstärkeren Lerngruppen. Kontrasteffekte scheinen dabei einen stärkeren Einfluss auf das akademische Selbstkonzept zu haben als Assimilationseffekte, wodurch insgesamt ein negativer Zusammenhang zwischen der Leistungsstärke der Lerngruppe und dem individuellem Selbstkonzept besteht (Marsh et al., 2000; Trautwein et al., 2006).

b) Soziale Unterstützung durch Lehrkräfte und Mitschülerinnen und Mitschüler

Neben sozialen Vergleichen beeinflussen auch Merkmale der sozialen Unterstützung durch Lehrkräfte und Mitschülerinnen und Mitschüler die schulische Motivation (Eccles, 2004; Patrick, Ryan & Kaplan, 2007; Wang & Eccles, 2013). Dabei lassen sich lernbezogene und emotionale Formen der sozialen Unterstützung unterscheiden (Song, Bong, Lee & Kim, 2014; Wentzel, 1998). Lernbezogene Unterstützungen umfassen Hilfestellungen durch Lehrkräfte und Mitschülerinnen und Mitschüler bei Unterrichtsaufgaben und Hausaufgaben. Dazu zählen auch kooperative Lernformen zwischen Schülerinnen und Schülern wie zum Beispiel Peer-Tutoring und die Bildung von Lerngruppen (Song et al., 2014; Wentzel et al., 2010; Wilkinson, 2002). Die emotionale Unterstützung umfasst positive Beziehungen zwischen Schülerinnen und Schülern und Lehrkräften, welche zu Gefühlen der Verbundenheit und des Angenommenseins führen (Patrick et al., 2007; Urdan & Schoenfelder, 2006; Wentzel, 1998). In Lerngruppen mit einem hohen Zusammengehörigkeitsgefühl erhalten Schülerinnen und Schüler mehr Unterstützung durch Mitschülerinnen und Mitschüler sowie durch Lehrkräfte und fühlen sich sicherer, aktiv am Unterricht teilzunehmen (Baumeister & Leary, 1995; Deci et al., 1991).

In dieser Dissertation wird die wahrgenommene emotionale Unterstützung der Schülerinnen und Schüler mit SPF in Förder- und Grundschulen durch ihre Mitschülerinnen und Mitschüler als Prädiktor für akademisches Selbstkonzept und Lernfreude betrachtet. In die Wahrnehmung der emotionalen Unterstützung fließen neben Interaktionen mit Mitschülerinnen und Mitschülern auch von der Lerngruppe geteilte Normen, Leistungserwartungen und Wertvorstellungen ein, die durch die Unterrichtsgestaltung und Schulkultur beeinflusst werden (Connell & Wellborn, 1991; Russell, 2012). Deshalb ist neben der individuellen Wahrnehmung der Schülerinnen und Schüler auch die geteilte Wahrnehmung der Lerngruppe im Hinblick auf die emotionale Unterstützung bedeutsam für die individuelle schulische Motivation (Russell, 2012). Anderman (2002) konnte zeigen, dass sich sowohl die individuelle Wahrnehmung als auch die auf Schulebene aggregierte emotionale Verbundenheit zur Schule auf Wohlbefinden, soziale Interaktionen beim Lernen und schulische Probleme auswirken und dadurch indirekt die schulische Motivation der Schülerinnen und Schüler beeinflussen.

Die emotionale Unterstützung in der Lerngruppe kann sich über verschiedene Vermittlungswege auf die schulische Motivation von Kindern und Jugendlichen auswirken (Forman, 1988; Wentzel et al., 2010). Auf der einen Seite führt eine hohe emotionale Unterstützung dazu, dass Schülerinnen und Schüler sich aktiv am Unterricht beteiligen können, ohne sich Sorgen

darüber zu machen, von Mitschülerinnen und Mitschüler verspottet zu werden (Furrer & Skinner, 2003; Patrick et al., 2007). Eine solche Lernumgebung begünstigt die intrinsische Motivation, das schulische Engagement und die Erfolgserwartungen der Schülerinnen und Schüler (Baumeister & Leary, 1995; Legault, Green-Demers & Pelletier, 2006; Skinner & Belmont, 1993). Somit bildet eine hohe soziale Integration eine Ressource für positive Einstellungen zum Lernen (Furrer & Skinner, 2003). Auf der anderen Seite wirkt die wahrgenommene soziale Unterstützung und Integration auch als Puffer gegen belastende Situationen im Unterricht (Demaray & Malecki, 2002; Furrer & Skinner, 2003; Wentzel, 1998). Schülerinnen und Schüler, die sich in einer Lerngruppe angenommen und unterstützt fühlen, können ihre schulische Motivation auch in Prüfungssituationen oder bei Misserfolgen besser aufrechterhalten (Baumeister & Leary, 1995; Deci & Ryan, 1985).

2.5.4.3 Argumente für höhere schulische Motivation in Förderschulen

Da sich Förderschulen durch reduzierte und weniger anspruchsvolle Lehrpläne und Unterrichtsaufgaben auszeichnen (Wocken, 2005), könnten Schülerinnen und Schüler mit SPF in Förderschulen häufiger schulische Erfolge erleben als in allgemeinen Schulen (Schumann, 2007; Tent et al., 1991). Diese Passung zwischen den Lernvoraussetzungen der Schülerinnen und Schüler und den eingesetzten Lehrmaterialien und Lernaufgaben trägt möglicherweise zur höheren schulischen Motivation von Kindern und Jugendlichen mit SPF bei (Eccles, 2004; Vaughn et al., 1996). So berichten Schülerinnen und Schüler höhere akademische Selbstkonzepte, wenn sie regelmäßig positive Leistungsrückmeldungen von ihren Lehrkräften erhalten (Elbaum, 2002; Pearl & Bay, 1999). Des Weiteren würden in Förderschulen häufiger individualisierte Leistungsrückmeldungen gegeben, die sich gemeinsam mit dem geringeren Leistungsdruck positiv auf das akademische Selbstkonzept und die Lernfreude der Kinder und Jugendlichen mit SPF auswirken könnten (Chapman, 1988; Markussen, 2004; Peetsma et al., 2001).

Außerdem lassen sich Kontrasteffekte als Argument für höhere motivationale Ausprägungen von Kindern und Jugendlichen mit SPF in Förderschulen anführen (Tracey & Marsh, 2000). In Förderschulen vergleichen Kinder und Jugendliche ihre Leistungen mit Mitschülerinnen und Mitschülern, die ähnliche Lernschwierigkeiten haben wie sie (Chapman, 1988). Außerdem ist die Leistungsstärke in Förderschulen geringer als in allgemeinen Schulen. Der soziale Vergleich mit leistungsschwächeren Mitschülerinnen und Mitschülern sollte zu höheren

akademischen Selbstkonzepten und zu einer höheren Lernfreude in Förderschulen führen (Tracey & Marsh, 2000). Soziale Vergleiche mit leistungsschwächeren Schülerinnen und Schülern verringern darüber hinaus schulischen Stress in Prüfungssituationen und könnten auch dadurch die schulische Motivation von Kindern und Jugendlichen mit SPF erhöhen (Dijkstra et al., 2008; Renick & Harter, 1989). Im Gegensatz dazu vergleichen Schülerinnen und Schüler mit SPF in allgemeinen Schulen ihre schulischen Kompetenzen mit leistungsstärkeren Mitschülerinnen und Mitschülern ohne SPF. Diese sozialen Vergleiche dürften dazu führen, dass Schülerinnen und Schüler mit SPF in allgemeinen Schulen niedrigere akademische Selbstkonzepte ausbilden als Förderschülerinnen und Förderschüler (Tracey & Marsh, 2000).

2.5.4.4 Argumente für eine höhere schulische Motivation in allgemeinen Schulen

Aber auch die Lernumgebung in allgemeinen Schulen könnte sich positiv auf die schulische Motivation von Kindern und Jugendlichen mit SPF auswirken. Befürworterinnen und Befürworter einer inklusiven Beschulung weisen darauf hin, dass Schülerinnen und Schüler mit SPF in allgemeinen Schulen von Lehrkräften stärker als fähige Lerner wahrgenommen werden (Bakker & Bosman, 2003). In allgemeinen Schulen würden Lehrkräfte höhere Leistungserwartungen und Fähigkeitseinschätzungen für alle Schülerinnen und Schüler haben, die sich in einem stärkeren Fokus auf den akademischen Fortschritt im Unterricht niederschlagen (Diamond et al., 2004; Hornstra, Denessen, Bakker, Bergh & Voeten, 2010). Positive Fähigkeitseinschätzungen der Lehrkräfte könnten von Kindern und Jugendlichen mit SPF übernommen werden und zu einer höheren schulischen Motivation in allgemeinen Schulen beitragen.

Ein weiteres Argument für eine höhere schulische Motivation von Kindern und Jugendlichen mit SPF in allgemeinen Schulen sind Assimilationseffekte. Allgemeine Schulen würden hinsichtlich der schulischen Leistungen ihrer Schülerschaft positiver bewertet als Förderschulen (Elbaum, 2002; Pfahl, 2011, 2012). Diese positive Bewertung könnten Schülerinnen und Schüler mit SPF in ihren Selbsteinschätzungen übernehmen, was zu einem höheren akademischen Selbstkonzept beitragen kann (Marsh et al., 2000; Tracey & Marsh, 2000). Zusätzlich könnte sich die schulische Motivation der Schülergruppe mit SPF dadurch erhöhen, dass sie positive Lernvorbilder in ihrer Lerngruppe vorfinden und diese als Ansporn für ihr schulisches Engagement nutzen (Corcoran et al., 2011; Wheeler & Suls, 2005).

In Förderschulen hingegen könne sich der hohe Anteil an Schülerinnen und Schülern mit problematischen Lernbiografien und Lernschwierigkeiten negativ auf die schulische Motivation auswirken. So würde die Schülerschaft in Förderschulen als leistungsschwach etikettiert und könnte dadurch von Stigmatisierungserfahrungen betroffen sein (Tracey & Marsh, 2000; Elbaum, 2002). Die Unterrichtsgestaltung in Förderschulen verfolge einen kompensatorischen Förderansatz, der die Aufmerksamkeit auf Defizite der Schülerinnen und Schüler lenkt (Wocken, 2005). Darüber hinaus sind der Besuch einer Förderschule und ein Förderschulabschluss mit geringen beruflichen und ausbildungsbezogene Karrierechancen verbunden (Pfahl, 2011; Schumann, 2007). Die Zuschreibung niedriger Fähigkeiten und Lernpotenziale durch schulische Akteure könne von den betroffenen Schülerinnen und Schülern wahrgenommen und internalisiert werden und zu kollektiv geteilten geringen Selbsteinschätzungen und Erfolgserwartungen führen (Knigge, 2017; Pfahl, 2011). Das kann zur Resignation von Förderschülerinnen und Förderschülern führen und ihre Anstrengungsbereitschaft und schulische Motivation reduzieren (*Marienthal-Effekt*; Jahoda, Lazarsfeld & Zeisel, 1975; s. a. Cooney, Jahoda, Gumley & Knott, 2006).

2.5.4.5 Empirische Befunde zum Einfluss der Beschulungsart auf die schulische Motivation von Kindern mit SPF

Die Mehrzahl der Studien zum Einfluss der Beschulungsart auf die schulische Motivation von Schülerinnen und Schülern mit SPF fokussiert auf das akademische Selbstkonzept. Dabei zeigten sich überwiegend Vorteile zugunsten von Schülerinnen und Schülern in Förderschulen beziehungsweise speziellen Förderklassen (Allodi, 2000; Bear, Minke & Manning, 2002; Chapman, 1988; Crabtree & Meredith, 2000; Renick & Harter, 1989; Tracey & Marsh, 2000). In einer aktuellen Studie fanden Szumski und Karwowski (2015), dass Grundschulkinder in Förderschulen höhere akademische Selbstkonzepte berichteten als Kinder mit SPF in allgemeinen Schulen.

Bear et al. (2002) schlussfolgerten in ihrer Metaanalyse ebenfalls, dass separate Lernumgebungen für Schülerinnen und Schüler mit SPF mit höheren akademischen Selbstkonzepten einhergehen als eine gemeinsame Beschulung von Kindern und Jugendlichen mit und ohne SPF. In ihrer Metaanalyse fassten sie die Ergebnisse von 61 Studien zusammen, die verschiedene Lernumgebungen (Förderschulen, Förderklassen, allgemeine Schulen und separate Lerngruppen für Kinder mit SPF in umgrenzten Fächern oder Zeiträumen) und verschiedene

Altersgruppen von Schülerinnen und Schülern mit SPF (vom Kindergarten bis zur Sekundarstufe) umfassten. Die Effekte zugunsten höherer akademischer Selbstkonzepte für Kinder und Jugendliche mit SPF in separaten Lerngruppen reichten von $d = 0.49$ (Cohen, 1992) im Bereich Schreiben, über $d = 0.66$ in Mathematik bis $d = 0.79$ im Lesen (Bear et al., 2002). Jedoch variierten diese Effekte in Abhängigkeit vom spezifischen Fördersetting. Diese Befunde geben Hinweise darauf, dass neben der Beschulungsart auch Merkmale der Lernumgebung (z. B. die Leistungsstärke der Lerngruppe) berücksichtigt werden sollten.

Im Gegensatz zu den Befunden von Bear et al. (2002) stellte Elbaum (2002) in ihrer Metaanalyse keinen bedeutsamen Zusammenhang zwischen der Beschulungsart und dem akademischen Selbstkonzept von Schülerinnen und Schülern mit SPF fest. Außerdem fand sie keine Hinweise darauf, dass Schülerinnen und Schüler in Förderschulen stärker stigmatisiert werden als Schülerinnen und Schüler mit SPF in allgemeinen Schulen. Jedoch blieben in der Studie individuelle Unterschiede in den Lernvoraussetzungen und dem soziokulturellen Hintergrund der Schülerinnen und Schüler mit SPF in Förderschulen und allgemeinen Schulen unberücksichtigt. Somit könnten individuelle Unterschiede zwischen den Schülergruppen den Effekt der Beschulungsart maskiert haben (Dijkstra et al., 2008).

Weiterhin können sich Schularteffekte auf die schulische Motivation in Abhängigkeit von der betrachteten Altersgruppe unterscheiden. Während Kinder zu Beginn der schulischen Laufbahn ihre eigenen Fähigkeiten stärker überschätzen, berichten sie am Ende der Grundschule realistischere und nach Schulfach differenzierte Selbsteinschätzungen (Helmke & Weinert, 1997; Shavelson et al., 1976). So fanden Butler und Marinov-Glasman (1994) für Drittklässler keine Unterschiede im akademischen Selbstkonzept zwischen Förderschülerinnen und -schülern und vergleichbaren Kindern in allgemeinen Schulen. Erst zu Beginn der 5. Klasse berichteten Kinder in Förderschulen höhere akademische Selbstkonzepte als leistungsschwache Kinder in allgemeinen Schulen. Dieses Ergebnis wurde dahingehend interpretiert, dass Fischteicheneffekte für die Ausbildung akademischer Selbstkonzepte relevant werden, wenn Kinder verstärkt soziale Vergleiche für ihre Selbsteinschätzungen heranziehen und sich ihre Fähigkeitseinschätzungen stärker zwischen unterschiedlichen Fächern ausdifferenzieren (Butler & Marinov-Glasman, 1994).

Peetsma et al. (2001) analysierten die Entwicklung der schulischen Motivation von Kindern mit SPF in der Primarstufe. Dabei wurden Kinder mit SPF in Förderschulen mit Kindern in allgemeinen Schulen verglichen. Die Vergleichsgruppe bestand aus Schülerinnen und Schülern ohne SPF, die ähnliche Lernvoraussetzungen (u. a. schulische Leistungen und

Selbstbild) und Merkmale der sozialen und kulturellen Herkunft (u. a. sprachliche Fähigkeiten, sozioökonomischer Status) wie die Schülergruppe mit SPF aufwiesen. Nach zwei Jahren schätzten Lehrkräfte die schulische Motivation von Kindern in Förderschulen im Vergleich zu Kindern in allgemeinen Schulen signifikant höher ein. Darüber hinaus verringerte sich die schulische Motivation von Kindern in allgemeinen Schulen über die ersten zwei Jahre der Untersuchung stärker als die von Kindern in Förderschulen. Nach vier Jahren zeigten sich jedoch keine bedeutsamen motivationalen Unterschiede zwischen den Schülergruppen mehr.

Der Forschungsstand im deutschsprachigen Raum zu motivationalen Unterschieden zwischen Kindern und Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen ist ebenfalls nicht eindeutig. Während frühe Studien Vorteile zugunsten von Förderschulen berichteten (Haeberlin, 1991; Rheinberg & Entstrup, 1977; Tent et al., 1991), sind die Ergebnisse neuerer Studien gemischt (Nusser & Wolter, 2016; Rossmann, Klicpera Gasteiger, Gebhardt, Roloff & Weindl, 2011; Schwinger et al., 2014; Wild et al., 2015). Beispielsweise fanden Rheinsberg und Enstrup (1977) in Einklang mit dem postulierten Fischteicheffekt, dass Kinder mit Lernbeeinträchtigungen in Förderschulen positivere Selbsteinschätzungen und geringere Testängstlichkeit aufwiesen als Kinder mit Lernbeeinträchtigungen in allgemeinen Schulen. Im Gegensatz dazu zeigten sich in einer neueren Untersuchung von Rossmann et al. (2011) keine bedeutsamen Unterschiede im akademischen Selbstkonzept und in der Testängstlichkeit zwischen Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen.

Auch erste Ergebnisse der bereits vorgestellten BiLieF-Studie (s. Kapitel 2.5.3.3) zeigen keine bedeutsamen motivationalen Unterschiede zwischen Kindern mit dem Förderschwerpunkt *Lernen* in Förderschulen und allgemeinen Schulen. So fanden sich in den querschnittlichen Analysen bisher lediglich etwas höhere Selbstwirksamkeitserwartungen für Kinder mit SPF in Förderschulen. Diese Effekte waren aber sehr klein, so dass ihre praktische Relevanz gering ist (Wild et al., 2015). In den weiteren betrachteten motivationalen Merkmale (Interesse, Anstrengungsbereitschaft, Zielorientierungen) unterschieden sich Kinder mit SPF in Förderschulen und allgemeinen Schulen nicht bedeutsam voneinander (Wild et al., 2015). Allerdings berichteten Kinder in Förderschulen ein signifikant höheres Stigmatisierungsempfinden als Kinder mit SPF in allgemeinen Schulen. Dieser Befund entspricht der Annahme von Assimilationsprozessen (Schwinger et al., 2014; s. a. Pfahl, 2011).

Hinweise für Kontrasteffekte bei der Ausbildung des akademischen Selbstkonzepts liefert eine aktuelle Untersuchung von Nusser und Wolter (2016). In dieser Studie verglichen die Autorinnen den Einfluss der mittleren Leistungsstärke einer Klasse auf das akademische

Selbstkonzept zwischen Förderschulen und allgemeinen Schulen. Insgesamt wurden 735 Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* (davon n = 148 Schülerinnen und Schüler in allgemeinen Schulen) in den Analysen berücksichtigt. Dabei zeigte sich nur für integrativ beziehungsweise inklusiv beschulte Kinder mit SPF ein negativer Zusammenhang zwischen der mittleren Leistungsstärke der Klasse und dem individuellen akademischen Selbstkonzept. In Förderschulen konnte unter Berücksichtigung individueller Leistungsunterschiede kein Fischteicheffekt nachgewiesen werden. Dieses Ergebnis interpretieren Nusser und Wolter (2016) so, dass individuell ausgerichtete Feedbackstrukturen in Förderschulen dazu beitragen können, dass soziale Vergleiche für Kinder und Jugendliche weniger stark hervortreten als in allgemeinen Schulen.

Insgesamt kann jedoch keine abschließende Aussage über den Einfluss der Beschulungsart auf die schulische Motivation von Schülerinnen und Schülern mit SPF getroffen werden. Es wurden nur wenige Studien durchgeführt, die die mittlere Leistungsstärke der Klassen oder die soziale Integration und Unterstützung als Erklärungsansätze für Schulartunterschiede einbeziehen. Darüber hinaus wird in der Mehrheit der Studien ausschließlich das akademische Selbstkonzept als Merkmal der motivationalen Entwicklung betrachtet, während zu intrinsischen Merkmalen der schulischen Motivation (z. B. Lernfreude) nur wenige Studien vorliegen (Fulk, Brigham & Lohman, 1998; Grolnick & Ryan, 1990; Peetsma et al., 2001). Außerdem werden in den meisten Studien Kinder und Jugendliche mit dem Förderschwerpunkt *Lernen* untersucht. Dabei bleibt offen, ob sich diese Ergebnisse auch auf Schülerinnen und Schüler mit den Förderschwerpunkten *Sprache* und *emotionale und soziale Entwicklung* übertragen lassen.

3

Forschungsfragen und Datengrundlage der Dissertation

3. Forschungsfragen und Datengrundlage der Dissertation

3.1 Forschungsstand und Forschungsbedarf zum Zusammenhang zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit SPF

Im Jahr 2016 wurden Schülerinnen und Schüler mit SPF in Deutschland noch überwiegend in Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2016). Im Zuge der Inklusionsbemühungen im deutschen Schulsystem wird jedoch zunehmend eine gemeinsame Beschulung dieser Schülergruppe in allgemeinen Schulen angestrebt (Prengel, 2013; Wrase, 2016). Dabei stellt sich aus Sicht der empirischen Bildungsforschung die Frage, in welchem Zusammenhang die Beschulungsart mit der schulischen Entwicklung von Kindern und Jugendlichen mit SPF steht.

Wie frühere Forschungsarbeiten vielfach belegten, stellen Schularten differenzielle Lern- und Entwicklungschancen für Kinder und Jugendliche bereit (Baumert et al., 2006; Nikolova, 2011). Diese Unterschiede zwischen Schularten lassen sich unter anderem auf institutionelle und kompositionelle Merkmale zurückführen (Baumert et al., 2006). Unter institutionellen Merkmalen werden unterschiedliche pädagogisch-didaktische Unterrichtstraditionen der Schularten verstanden, die sich in den Lehrplänen, der Ausbildung und Qualifikation der Lehrkräfte und der Organisation des Unterrichts widerspiegeln (Gamoran, 1986; Köller & Baumert, 2012). Zum Beispiel wird angenommen, dass die Lehrpläne und die Unterrichtsgestaltung in allgemeinen Schulen anspruchsvoller und abwechslungsreicher sind als in Förderschulen (Hocutt, 1996; Schumann, 2007; Wocken, 2000, 2005). Förderschulen hingegen könnten einen Schonraum darstellen, von dem Schülerinnen und Schüler mit SPF durch eine auf ihre Lernvoraussetzungen angepasste Unterrichtsgestaltung in ihrer schulischen Entwicklung profitieren (Schumann, 2007; Tent et al., 1991; Vaughn et al., 1996; Wocken, 1996).

Auch kompositionelle Merkmale, wie die leistungsbezogene und soziale Zusammensetzung der Schülerschaft, können sich auf die individuelle schulische Entwicklung auswirken (Dumont et al., 2013; Harris, 2010; Justice et al., 2014; Thrupp et al., 2002). Möglicherweise ist das Lernklima in Förderschulen durch den hohen Anteil an Kindern und Jugendlichen mit geringen sozioökonomischen und kulturellen Ressourcen sowie problematischen Lernbiogra-

fien weniger günstig als in allgemeinen Schulen (Bos et al., 2010; Wocken, 2005). Darüber hinaus passen Lehrkräfte ihr Anforderungsniveau und ihre Leistungserwartungen an die Leistungsstärke der Klasse an (Diamond et al., 2004; Harris, 2010; Rist, 2000; Weinert & Helmke, 1997). So könnten Lehrkräfte in Förderschulen das Anforderungsniveau absenken und niedrigere Erwartungen hinsichtlich der leistungsbezogenen und motivationalen Entwicklung der Schülerinnen und Schüler mit SPF bilden, die sich auf ihre schulische Entwicklung auswirken.

Der Forschungsstand zum Zusammenhang zwischen Beschulungsart und schulischen Outcomes erlaubt noch kein abschließendes Urteil darüber, ob Schülerinnen und Schüler mit SPF von einer Beschulung in Förderschulen oder in allgemeinen Schulen stärker profitieren (Grosche & Vock, im Druck). Als zentrale Merkmale der schulischen Entwicklung werden in dieser Dissertation die schulischen Kompetenzen und die schulische Motivation in den Blick genommen. Für die schulischen Kompetenzen von Schülerinnen und Schülern mit SPF fanden sich in der Mehrzahl der Studien Vorteile zugunsten allgemeiner Schulen (Baker et al., 1994; Carlberg & Kavale, 1980; Peetsma et al., 2001; Ruijs & Peetsma, 2009; Wocken, 2000, 2005). Die Größe dieser Effekte variiert dabei von geringen (Baker et al., 1994; Carlberg & Kavale, 1980, Stranghöner et al., 2017) bis mittelstarken Leistungsvorsprüngen (Wang & Baker, 1985). In einzelnen Studien fanden sich allerdings auch höhere schulische Leistungen für Schülerinnen und Schüler mit SPF in Förderschulen beziehungsweise Förderklassen (Cole, Waldron & Majd, 2004; Harrington, 1997; Lehmann & Hoffmann, 2009; Lindsay, 2007; Rogers & Thiery, 2003). Für Merkmale der schulischen Motivation weist die Mehrzahl empirischer Studien auf Vorteile zugunsten von Kindern und Jugendlichen mit SPF in Förderschulen beziehungsweise separaten Förderklassen hin (Bear et al., 2002; Chapman, 1988; Renick & Harter, 1989; Szumski & Karwowski, 2015; Tent et al., 1991; Tracey & Marsh, 2000). Jedoch variiert die Stärke der Schularteffekte zwischen diesen Studien (Bear et al., 2002; Elbaum, 2002). Andere Studien fanden keine bedeutsamen Unterschiede in der schulischen Motivation zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen (Rossmann et al., 2011; Wild et al., 2015).

Wesentliche Limitationen der vorliegenden Studien bestehen darin, dass sie nur selten standardisierte Leistungstests verwendet und adäquate Vergleichsgruppen einbezogen haben. So werden in manchen Untersuchungen Schülerinnen und Schüler ohne SPF als Vergleichsgruppe herangezogen und deren Entwicklung mit der von Schülerinnen und Schülern an Förderschulen verglichen (Wocken, 2000, 2005). Außerdem wurden in den Untersuchungen

im deutschsprachigen Raum bislang nur regional begrenzte Stichproben untersucht. Bisherige Studien fokussierten zudem auf Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen*, während andere Förderschwerpunkte bislang kaum Beachtung fanden. Eine offene Forschungsfrage besteht daher darin, ob sich für verschiedene Förderschwerpunkte ähnliche Zusammenhänge zwischen der Beschulungsart und den schulischen Kompetenzen beziehungsweise Merkmalen der schulischen Motivation zeigen.

Eine weitere Limitation bisheriger Studien liegt darin, dass nur wenige Hintergrundmerkmale der Schülerinnen und Schüler mit SPF in den Analysen berücksichtigt wurden. Dadurch könnten die festgestellten Schularteffekte teilweise auch auf unterschiedliche Eingangs- und Lernvoraussetzungen von Kindern und Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen zurückzuführen sein (Zigmond, 2003). Deshalb muss zunächst geprüft werden, ob sich Schülerinnen und Schüler mit SPF in ihren individuellen Lernvoraussetzungen und soziokulturellen Merkmalen zwischen Förderschulen und allgemeinen Schulen unterscheiden. Die Auswertungen empirischer Studien und amtlicher Schulstatistiken weisen darauf hin, dass Kinder und Jugendliche mit SPF in Förderschulen über weniger soziale und kulturelle Ressourcen verfügen als Schülerinnen und Schüler in allgemeinen Schulen (Autorengruppe Bildungsberichterstattung, 2012; Klein, 2001; Wocken, 2000, 2005). Weiterhin gibt es Hinweise darauf, dass Kinder und Jugendliche aus zugewanderten Familien insbesondere in Förderschulen mit dem Schwerpunkt *Lernen* überrepräsentiert sind (Dietze, 2011; Kemper & Weishaupt, 2011; Kornmann, 2006; Kronig, 2003; Powell & Wagner, 2002). In der Mehrzahl dieser Studien wurden jedoch Förderschülerinnen und -schüler mit Kindern und Jugendlichen ohne SPF in allgemeinen Schulen verglichen. Es besteht Forschungsbedarf zur Frage, ob sich auch Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen in diesen Merkmalen unterscheiden.

Unterschiedliche Lernvoraussetzungen und soziokulturelle Hintergründe von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen sollten zum einen berücksichtigt werden, um präzisere Schätzungen von Schularteffekten zu ermöglichen. Zum anderen können sie zu schulartspezifischen Schülerzusammensetzungen beitragen. Empirische Studien zeigen, dass die fähigkeitsbezogene und auch soziokulturelle Klassenzusammensetzung mit individuellen Schulleistungen und Merkmalen der schulischen Motivation von Kindern und Jugendlichen mit SPF zusammenhängen (Justice et al., 2014; Nusser & Wolter, 2016; Rauer & Schuck, 2007). Forschungsbedarf besteht vor allem dahingehend, welche Merkmale der Lernumgebungen zu motivationalen Unterschieden zwischen Schülerinnen und

Schülern mit SPF beitragen. Insbesondere im deutschsprachigen Raum wurden nur wenige Studien durchgeführt, die Merkmale der Klassenkomposition als mögliche Erklärungen für Schulartunterschiede einbeziehen (z. B. Nusser & Wolter, 2016).

3.2 Datengrundlage der Dissertation

Die vorliegende Dissertation basiert auf Daten des IQB-Ländervergleichs 2011 für die Primarstufe (Stanat et al., 2012; Stanat et al., 2014). Dieses querschnittlich angelegte Large-Scale Assessment im Bildungsbereich diente der Überprüfung des Erreichens der Bildungsstandards am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Im IQB-Ländervergleich 2011 wurde eine auf Bundes- und Länderebene repräsentative Stichprobe von insgesamt 27 081 Kindern aus 1 349 Schulen getestet. Darunter befanden sich auch 1 195 Schülerinnen und Schüler mit einem diagnostizierten SPF im Bereich *Lernen*, *Sprache* oder *emotionale und soziale Entwicklung*. Die Mehrheit dieser Schülergruppe besuchte eine allgemeine Schule ($n = 762$ Kinder aus 485 Grundschulen), während 433 Kinder mit SPF in 51 Förderschulen unterrichtet wurden. Um nach Förderschwerpunkt differenzierte Zusammenhänge zwischen Beschulungsart und schulischen Outcomes zu analysieren, wurden Kinder mit mehreren Förderschwerpunkten beziehungsweise Doppeldiagnosen nicht in die Analysestichprobe aufgenommen. Die meisten Schülerinnen und Schüler in der vorliegenden Stichprobe hatten einen SPF mit dem Förderschwerpunkt *Lernen* ($n = 614$), gefolgt von Kindern mit dem Förderschwerpunkt *Sprache* ($n = 326$) und Kindern mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* ($n = 255$). Aufgrund der geringen Anzahl von Kindern mit Förderschwerpunkt *emotionale und soziale Entwicklung* in Förderschulen ($n = 20$) konnte für diese Schülergruppe keine förderschwerpunktspezifische Auswertung von möglichen Schulartunterschieden vorgenommen werden.

In Einklang mit anderen Large-Scale Assessments im Bildungsbereich wurden nur Schülerinnen und Schüler in die Analysestichprobe einbezogen, die an der Kompetenztestung teilnehmen konnten und prinzipiell in der Lage war, die Fragebögen und Kompetenztests eigenständig im Rahmen der festgelegten Testbedingungen zu bearbeiten (Joncas & Foy, 2012; OECD, 2014). So wurden Schülerinnen und Schüler von der Testung ausgeschlossen, die eine dauerhafte körperliche Beeinträchtigung aufwiesen, die eine Teilnahme an den Leistungstests unmöglich machte. Auch nahmen Schülerinnen und Schüler nicht an der Testung teil, wenn

sie emotional oder geistig nicht in der Lage waren, den allgemeinen Anweisungen der Tests zu folgen, oder die durch die Testsituation in unzumutbarem Maße emotional belastet worden wären. Schließlich wurden Schülerinnen und Schüler nicht in die Analysen einbezogen, deren Muttersprache nicht Deutsch war und die weniger als ein Jahr in deutscher Sprache unterrichtet worden waren (Richter, Engelbert et al., 2012). Insgesamt wurden 41 Kinder mit SPF in den drei betrachteten Förderschwerpunkten (3.5%) aufgrund eines dieser Ausschlusskriterien nicht in die Analysen einbezogen.

Ziel von Large-Scale Assessments im Bildungsbereich ist es, schulische Kompetenzen (und andere schulische Outcomes) von Schülerinnen und Schülern zu erfassen und dadurch auch Hinweise für möglichen Steuerungsbedarf auf Ebene von Bildungssystemen bereitzustellen (Chudowsky & Pellegrino, 2003). Die in LSAs eingesetzten Testinstrumente wurden allerdings nicht für Kinder und Jugendliche mit SPF entwickelt und erprobt. Deshalb konnte nicht vorausgesetzt werden, dass die Testinstrumente und Fragebögen bei Schülerinnen und Schülern mit SPF die zugrundeliegenden Konstrukte (z. B. schulische Kompetenzen) in vergleichbarer Art und Weise erfassen wie bei Schülerinnen und Schülern ohne SPF. Diese Annahme ist jedoch notwendig, um eine gemeinsame Skala für beide Schülergruppen zu definieren und ihr Abschneiden darauf abzubilden. In nationalen Schulleistungsstudien wurde dies bislang selten empirisch geprüft (Heydrich et al., 2013). Zwei Ausnahmen sind die Zusatzerhebungen im Rahmen von PISA 2012 (Müller, Sälzer, Mang, Heine et al., 2014) und NEPS (Pohl et al., 2016; Südkamp et al., 2015). In beiden Studien zeigte sich, dass die Messeigenschaften der Kompetenztests zwischen der Schülergruppe mit SPF und der Schülergruppe ohne SPF nur eingeschränkt miteinander vergleichbar sind (Müller, Sälzer, Mang, Heine et al., 2014; Südkamp et al., 2015). Beide Studien betrachteten dabei ausschließlich Schülerinnen und Schüler mit SPF in Förderschulen. Die vorliegende Dissertation geht darüber hinaus und vergleicht Schülerinnen und Schüler mit SPF in allgemeinen Schulen mit Schülerinnen und Schülern mit SPF in Förderschulen.

Um valide Schulartvergleiche durchführen zu können, sollten die Testinstrumente auch zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen vergleichbare Messeigenschaften aufweisen. Ist das nicht der Fall, können die Testwerte beider Schülergruppen nicht vergleichbar interpretiert werden. Als Folge könnten festgestellte Kompetenzunterschiede zwischen den Schülergruppen nicht nur auf Schularteffekte, sondern auch auf unterschiedliche Messeigenschaften der Testinstrumente zurückzuführen sein (Lane & Leventhal, 2015). Eine offene Forschungsfrage ist, ob die in Schulleistungsstudien

eingesetzten Kompetenztests Fähigkeitsunterschiede bei Kindern mit SPF in beiden Schularten gleichermaßen reliabel und valide abbilden können.

3.3 Forschungsfragen und Teilstudien

Ziel der vorliegenden Dissertation ist es, den Zusammenhang zwischen Beschulungsart und schulischen Kompetenzen sowie schulischer Motivation von Kindern mit SPF in der Primarstufe zu untersuchen. Dabei werden Unterschiede in der Klassenzusammensetzung zwischen Förderschulen und allgemeinen Schulen als eine mögliche Erklärung für Schulartunterschiede analysiert. Außerdem soll geprüft werden, ob sich der Zusammenhang zwischen Beschulungsart und schulischen Outcomes in Abhängigkeit vom Förderschwerpunkt der Schülerinnen und Schüler mit SPF unterscheidet. Aus dem beschriebenen Forschungsstand und dem dargestellten Forschungsbedarf lassen sich vier Fragestellungen ableiten:

1. Sind die im IQB-Ländervergleich 2011 eingesetzten Kompetenztests geeignet, die schulischen Kompetenzen von Kindern mit SPF in Förder- und Grundschulen reliabel und valide zu erfassen?
2. Unterscheiden sich die schulischen Kompetenzen von Kindern mit SPF in Abhängigkeit von der Beschulungsart?
3. Können mögliche Kompetenzunterschiede zwischen Kindern mit SPF in Förder- und Grundschulen auf individuelle Unterschiede in Eingangs- und Lernvoraussetzungen sowie schulartspezifischen Klassenzusammensetzungen zurückgeführt werden?
4. Unterscheiden sich Kinder mit SPF in Förder- und Grundschulen in ihrer schulischen Motivation?

Diese vier Fragestellungen wurden in vier Teilstudien bearbeitet:

Teilstudie 1: Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

In der ersten Studie wurde – als grundlegende Voraussetzung für alle weiteren Analysen – geprüft, ob die im IQB-Ländervergleich 2011 eingesetzten Kompetenztests für Kinder mit SPF in Förder- und Grundschulen vergleichbar gut geeignet sind, deren schulischen Kompetenzen zu erfassen. Darüber hinaus untersuchten wir, ob die Messeigenschaften der Kompetenztests zwischen den Schülergruppen mit und ohne SPF vergleichbar sind (Fragestellung 1).

Zu diesem Zweck verglichen wir Aspekte der Testgüte der Kompetenztests zwischen den drei Schülergruppen (Kinder ohne SPF, Kinder mit SPF in Grundschulen, Kinder mit SPF in Förderschulen). Zur Bestimmung der Kompetenzwerte der Schülerinnen und Schüler wurden deren Antworten in den Kompetenztests anhand eines eindimensionalen *Item Response Theory* (IRT)-Modells (Rasch-Modell, vgl. Embretson & Reise, 2000) skaliert. Die Vergleichbarkeit der Testwerte für Schülerinnen und Schüler mit SPF prüften wir in vier Schritten: Zunächst wurde der Anteil fehlender Werte analysiert. Diese Analysen geben Hinweise darauf, ob Kinder mit SPF die Testaufgaben unter den gegebenen Testbedingungen bewältigen können. Im zweiten Schritt prüften wir die Passung des Rasch-Modells. Dabei wurden die Testitems für jede betrachtete Schülergruppe separat skaliert und anschließend die Infitstatistiken miteinander verglichen. Im dritten Schritt wurden Analysen zu differenziellen Itemfunktionen (DIF, Holland & Thayer, 1988) durchgeführt. DIF liegt vor, wenn zum Beispiel Kinder mit SPF in Förderschulen und Kinder mit SPF in Grundschulen trotz gleichen Fähigkeitsniveaus ein Item mit einer unterschiedlichen Wahrscheinlichkeit lösen. Im vierten Schritt überprüften wir, ob die korrelativen Zusammenhänge der Testwerte im Lesen, Zuhören und in Mathematik untereinander sowie zu den verbalen und nonverbalen kognitiven Grundfähigkeiten in den Schülergruppen ähnlich ausgeprägt sind. Daraus können Hinweise auf die Vergleichbarkeit der externen Validität der Testwerte gewonnen werden.

Teilstudie 2: Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe

In der zweiten Teilstudie wurde betrachtet, ob sich die schulischen Kompetenzen von Kindern mit SPF in Förder- und Grundschulen unterscheiden (Fragestellung 2). Dabei differenzierten wir sowohl nach Kompetenzbereich beziehungsweise Fach (Deutsch Lesen, Deutsch Zuhören und Mathematik) als auch nach Förderschwerpunkt (*Lernen* vs. *Sprache*). Da Schülerinnen und Schüler nicht zufällig einer Schulart zugewiesen werden, könnten systematische Unterschiede in individuellen Lernvoraussetzungen zu Kompetenzunterschieden zwischen Kindern mit SPF in Förder- und Grundschulen beitragen. Um für diese möglichen Selektionseffekte zu kontrollieren, setzten wir verschiedene Propensity-Score-Matching-Verfahren ein. Im Rahmen dieser Verfahren werden nur Kinder mit ähnlichen Lernvoraussetzungen in Förder- und Grundschulen miteinander verglichen. Nach erfolgreichem Matching können gefundene Unterschiede in den schulischen Kompetenzen mit höherer Plausibilität auf die Schulart zurückgeführt werden. Wir nahmen an, dass nach dem Matching bedeutsame Leistungsvor-

sprünge für Kinder mit SPF in Grundschulen im Vergleich zu Kindern in Förderschulen festgestellt werden können.

Teilstudie 3: Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf

In Teilstudie 3 wurde geprüft, ob Merkmale der Klassenkomposition zu Kompetenzunterschieden von Kindern mit SPF in Förder- und Grundschulen beitragen (Fragestellung 3). Mit Hilfe von Mehrebenenmodellen (Raudenbush & Bryk, 2002) analysierten wir, ob die soziale und fähigkeitsbezogene Klassenzusammensetzung (Komposition) sowie die Schulart (Institution) über individuelle Schülermerkmale hinaus mit Unterschieden in den schulischen Kompetenzen von Kindern mit SPF zusammenhängen. Dabei spezifizierten wir Modelle, in die schrittweise Prädiktoren auf Individual- und Klassenebene aufgenommen wurden. Ziel dieses schrittweisen Vorgehens war es, zu prüfen, ob Merkmale der Klassenzusammensetzung auch nach Berücksichtigung von individuellen Prädiktoren Kompetenzunterschiede von Kindern mit SPF in Förder- und Grundschulen erklären. In Einklang mit dem Forschungsstand erwarteten wir, dass Unterschiede in der Klassenkomposition auch nach Kontrolle individueller Schülermerkmale mit Kompetenzunterschieden zwischen Kindern mit SPF in Förder- und Grundschulen zusammenhängen. Dabei sollte insbesondere die fähigkeitsbezogene Klassenzusammensetzung positiv mit den schulischen Kompetenzen von Kindern mit SPF zusammenhängen.

Teilstudie 4: Educational placement and achievement motivation of students with special educational needs

In der vierten Teilstudie stand die Frage im Vordergrund, ob Kinder mit SPF in Förderschulen oder Grundschulen eine höhere schulische Motivation berichten (Fragestellung 4). Darüber hinaus untersuchten wir, ob motivationale Unterschiede mit der Klassenzusammensetzung (mittlere Leistungsstärke und wahrgenommene soziale Unterstützung) in Förder- und Grundschulen zusammenhängen. Zu den betrachteten Merkmalen der schulischen Motivation gehörten das akademische Selbstkonzept und die Lernfreude der Kinder mit SPF. Außerdem wurde analysiert, ob sich diese Zusammenhänge zwischen Kindern mit dem Förderschwerpunkt *Lernen* und *Sprache* unterscheiden. Zur Beantwortung der Fragestellungen kamen ebenfalls Mehrebenenmodelle zum Einsatz. Wir vermuteten, dass Kinder mit SPF in Förderschulen nach Berücksichtigung individueller Leistungsunterschiede eine höhere schulische Motivation berichten als Kinder mit SPF in Grundschulen. In Einklang mit dem Fischteichf-

fekt (Big-Fish-Little-Pond-Effect; Marsh, 1987) sollten diese motivationalen Unterschiede auf die niedrigere durchschnittliche Leistung der Mitschülerinnen und Mitschüler in Förderschulen zurückführbar sein. Darüber hinaus nahmen wir an, dass auch die soziale Unterstützung durch Mitschülerinnen und Mitschüler positiv mit der schulischen Motivation zusammenhängen (Deci et al., 1991; Wentzel et al., 2010).

4

Teilstudien der Dissertation

4. Teilstudien der Dissertation

4.1 Teilstudie 1: Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

Abstract

Im vorliegenden Beitrag wurde geprüft, inwieweit die im IQB-Ländervergleich 2011 in der Primarstufe eingesetzten Kompetenztests in Deutsch und Mathematik dazu geeignet sind, auch die Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) adäquat zu erfassen. Dazu wurden der Anteil fehlender Werte, die Passung des Raschmodells, differenzielle Itemfunktionen und Korrelationen der Testwerte zwischen den Schülergruppen mit SPF in Förder- und Grundschulen und Kindern ohne SPF analysiert. Die eingesetzten Tests scheinen die schulischen Kompetenzen von Kindern mit SPF reliabel und valide zu erfassen. Allerdings zeigten sich Einschränkungen hinsichtlich der Vergleichbarkeit der Messungen für Kinder in Förderschulen in Mathematik.

Schlagwörter: Sonderpädagogischer Förderbedarf, Testgüte, Schulische Kompetenzen, Primarstufe, Beschulungsart

Teilstudie 1 ist in folgender Publikation erschienen:

Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A. & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion*. 62. Beiheft der Zeitschrift für Pädagogik (S. 212–234). Weinheim: Beltz Juventa.

4.1.1 Einleitung

Large-Scale Assessments (LSAs) im Rahmen des Bildungsmonitorings zielen darauf ab, schulische Kompetenzen von Schülerinnen und Schülern zu messen und Unterschiede in den Erträgen von Bildungssystemen insgesamt sowie für verschiedene Gruppen von Lernenden sichtbar zu machen. Damit sollen die Studien Stärken und Schwächen auf der Systemebene identifizieren und Impulse für eine Verbesserung der Bedingungen schulischen Lernens geben (Chudowsky & Pellegrino, 2003). Diese Konzeptualisierung von LSAs setzt voraus, dass ihre Ergebnisse die jeweilige Zielpopulation und relevante Teilpopulationen adäquat repräsentieren (Olson & Goldstein, 1997).

Eine Schülergruppe, die in den letzten Jahren verstärkt in den Fokus der empirischen Bildungsforschung gerückt ist, ist die der Schülerinnen und Schüler, bei denen ein sonderpädagogischer Förderbedarf (SPF) diagnostiziert wurde. Diese Gruppe wird im Rahmen von nationalen und internationalen LSAs (z.B. Ländervergleich des IQB, PISA, IGLU) soweit wie möglich als Teil der Zielpopulation definiert⁵ und in die Stichprobenziehung eingeschlossen. Damit soll gewährleistet werden, dass die Schülerschaft, auf die sich das Monitoring bezieht, möglichst breit erfasst wird.

Die Einbeziehung von Schülerinnen und Schülern mit SPF im Bildungsmonitoring ist aber vor allem auch aufgrund des steigenden Anteils von Schülerinnen und Schülern mit SPF an Regelschulen des allgemeinbildenden Schulsystems relevant (Autorengruppe Bildungsberichterstattung, 2014). Die UN-Behindertenrechtskonvention (2008) weist darauf hin, dass empirische Daten bereitgestellt werden müssen, um die Entwicklung eines inklusiven Bildungssystems evaluieren zu können und der Bildungspolitik Anhaltspunkte für eine effektive Steuerung der Umsetzung dieser Reform zu liefern (Artikel 31, UN-BRK). Zur Umsetzung der auf Inklusion bezogenen rechtlichen Vorgaben ist es also wichtig, Schülerinnen und Schüler mit SPF in Schulleistungsstudien einzubeziehen (Thurlow, 2002). LSAs können als wichtiges Instrument zur Evaluation inklusiver Bemühungen im deutschen Bildungssystem genutzt werden (Wrase, 2015), um die schulische Entwicklung von Kindern mit SPF zu verfolgen (Abedi et al., 2011).

⁵ Vorab aus der Zielpopulation ausgeschlossen werden Schülerinnen und Schüler, die aufgrund sehr geringer Sprachkenntnisse (Beschulungsdauer in der Testsprache von weniger als einem Jahr) oder einer Beeinträchtigung nicht in der Lage wären, die Tests zu bearbeiten.

Aufgrund dieser Zielperspektiven wird zunehmend versucht, die Einbeziehung von Kindern und Jugendlichen mit SPF in nationalen und internationalen Schulleistungsstudien zu optimieren (Thurlow, 2010). Im Fokus nationaler Schulleistungsstudien stehen dabei vor allem Schülerinnen und Schüler mit den Förderschwerpunkten *Lernen*, *Sprache* sowie *emotionale und soziale Entwicklung*. Dabei handelt es sich – neben *geistiger Entwicklung* – um die drei größten Förderschwerpunkte in Deutschland (Autorengruppe Bildungsberichterstattung, 2014), wobei die Differenzierung zwischen den Förderschwerpunkten aufgrund von substanziellen Überschneidungen in Einzelfällen schwierig ist (Opp, Budnik & Fingerle, 2008; Ricking, 2005). Auch ist die Abgrenzung dieser drei Schülergruppen mit SPF von Schülerinnen und Schülern ohne SPF nur graduell möglich und scheint zum Teil durch soziale Bezugsnormen beeinflusst zu werden (Bos et al., 2010; Kornmann, 2006). Für diese Schülergruppen wird daher zunehmend diskutiert, auf eine separate Beschulung in Förderschulen vollständig zu verzichten und sie stattdessen in Regelschulen des allgemeinbildenden Schulsystems zu unterrichten (Autorengruppe Bildungsberichterstattung, 2014).

Die Kompetenztestung von Schülerinnen und Schülern mit SPF in LSAs wird kontrovers diskutiert (Smith & Douglas, 2014; Thurlow, 2002). Die primäre Funktion von LSAs besteht in der Evaluation von schulischen Outcomes auf der Systemebene. Impulse für die Unterrichtsentwicklung und die konkrete Ausgestaltung von Inklusion sind somit nicht Ziel von LSAs. Allerdings befürchten Kritiker, dass die für LSAs konzipierten Tests auch für diese Zwecke verwendet werden könnten (Smith & Douglas, 2014). Dies würde zu einer zu starken Betonung schulischer Leistungen bei der Evaluation inklusiver Bemühungen führen (Smith & Douglas, 2014). Darüber hinaus ermöglicht die überwiegend querschnittliche Ausrichtung von LSAs nicht, Leistungsentwicklungen von Schülerinnen und Schülern mit SPF abzubilden. Im Rahmen eines inklusiven Ansatzes wird argumentiert, dass weitere schulische Outcomes (z.B. soziale Kompetenzen, Motivation, Wohlbefinden) und Entwicklungsverläufe von Schülerinnen und Schülern mit SPF betrachtet werden sollten (Smith & Douglas, 2014). Auf der anderen Seite stellen LSAs eine wichtige Grundlage dar, um den Bildungserfolg der Schülergruppe mit SPF im Hinblick auf die weitere schulische und berufliche Ausbildung zu evaluieren. Zudem könnten hohe Leistungserwartungen an alle Schülerinnen und Schüler die Erwartungshaltung und Verantwortungsbereitschaft der Lehrkräfte erhöhen, wodurch auch die Schülergruppe mit SPF in ihrer Leistungsentwicklung profitieren könnte (Thurlow, 2002). Schließlich besteht ein hoher gesellschaftlicher Konsens darüber, dass Schülerinnen und Schülern mit SPF Zugang zu den gleichen schulischen Erfahrungen ermöglicht werden sollte, die auch Kinder ohne SPF erfahren (Pitoniak & Royer, 2001). LSAs ersetzen somit nicht die

Erfassung individueller Lernfortschritte in akademischen, motivationalen und sozialen Bereichen, sie können aber ein nützliches Instrument zur Evaluation des schulischen Abschneidens von Kindern mit SPF darstellen.

Um aus dem Abschneiden von Schülerinnen und Schülern mit SPF in LSAs Schlussfolgerungen für inklusive Bemühungen im deutschen Bildungssystem ableiten zu können, muss zunächst geprüft werden, inwieweit die derzeit eingesetzten Kompetenztests für diese Schülergruppe geeignet sind. Die Tests wurden ursprünglich für Schülerinnen und Schüler ohne SPF entwickelt und erprobt, sie werden jedoch auch zur Erfassung der Kompetenzen von Schülerinnen und Schülern mit einem Förderbedarf in den Bereichen *Lernen*, *Sprache* oder *emotionale und soziale Entwicklung* eingesetzt. Im Folgenden soll der Frage nachgegangen werden, inwieweit die Kompetenztests dafür geeignet sind, wobei die Prüfung sowohl für Kinder mit SPF an Förderschulen als auch für Kinder mit SPF an Grundschulen vorgenommen wird.

4.1.2 Methodische Überlegungen

Die zentrale Herausforderung bei der gemeinsamen Testung von Schülerinnen und Schülern mit und ohne SPF besteht darin, die schulischen Leistungen der gesamten Schülerschaft auf einer gemeinsamen Skala abzubilden und gleichzeitig für jede Schülergruppe eine reliable und valide Testung zu gewährleisten (Heydrich, Weinert, Nusser, Artelt & Carstensen, 2013). Damit verbunden ist die Frage der Äquivalenz der Testergebnisse von Heranwachsenden mit SPF und Heranwachsenden ohne SPF (Lane & Leventhal, 2015). Eine eingeschränkte Vergleichbarkeit der Testergebnisse könnte dazu führen, dass Unterschiede zwischen diesen beiden Schülergruppen im interessierenden Kompetenzkonstrukt (z.B. mathematische Kompetenz) durch die Testwertunterschiede nicht oder nur verzerrt abgebildet werden.

Um die Testgüte bzw. Validität schulischer Leistungstests zu beurteilen, muss geprüft werden, ob die Testwerte die zugrundeliegenden Fähigkeiten adäquat erfassen und ob die Interpretation der Testwerte für verschiedene Schülergruppen vergleichbar ist (Messick, 1995). In der Literatur werden verschiedene Aspekte der Validität diskutiert (Messick, 1995). In der vorliegenden Untersuchung wird die *Generalisierbarkeit* der Interpretation von Testwerten aus Schulleistungserhebungen auf die Teilpopulation der Schülerinnen und Schüler mit SPF in den Bereichen *Lernen*, *Sprache* sowie *emotionale und soziale Entwicklung*

in den Blick genommen. Es wird geprüft, ob die Gütekriterien der Kompetenztests für Kinder mit SPF in Förder- und Grundschulen in vergleichbarer Weise wie für Kinder ohne SPF erfüllt sind und inwieweit sich die Interpretation der Testwerte bei Kindern ohne SPF auf die Schülergruppe mit SPF übertragen lässt. Zur Beurteilung der Generalisierbarkeit werden strukturelle und externe Aspekte der Validität betrachtet. Die Vergleichbarkeit der Aussagekraft von Testwerten für verschiedene Schülergruppen als ein Merkmal der Testgüte bzw. Validität wird auch unter dem Stichwort der „Testfairness“ diskutiert (Cole & Zieky, 2001). Bleiben Aspekte der Testfairness ungeprüft, könnten beobachtete Gruppenunterschiede nicht allein auf Unterschiede zwischen Schülergruppen im zugrundeliegenden Konstrukt, sondern auch auf Merkmale der Testinstrumente zurückzuführen sein (Cole & Zieky, 2001). Die Validität der Ergebnisse bzw. deren Interpretation wäre also eingeschränkt.

Die angemessene Interpretation von Gruppenunterschieden in Testergebnissen setzt voraus, dass strukturelle Aspekte der Testverfahren zwischen Schülergruppen vergleichbar sind (Messick, 1995). Insbesondere sollten die interne Struktur der Leistungstests und die psychometrischen Eigenschaften der Testitems zwischen den Schülergruppen hinreichend ähnlich sein. Strukturelle Aspekte und psychometrische Eigenschaften der im Folgenden analysierten Kompetenztests wurden für Schülerinnen und Schüler ohne SPF bereits bestimmt und validiert (Böhme & Robitzsch, 2009; Winkelmann & Robitzsch, 2009). Diese Kennwerte dienen in der vorliegenden Untersuchung als Referenz zur Überprüfung der Testgüte für Kinder mit SPF in Förder- und Grundschulen. Anhand der bereits vorliegenden Auswertungsmodelle wird geprüft, ob die Kompetenztests in vergleichbarer Weise geeignet sind, die Kompetenzen von Schülerinnen und Schülern mit SPF an Förder- und Grundschulen zu erfassen wie für Schülerinnen und Schüler ohne SPF. Nur wenn gesichert ist, dass die Kompetenztests bei Kindern mit und bei Kindern ohne SPF eine ähnliche Struktur aufweisen, kann vorausgesetzt werden, dass die in LSAs eingesetzten Testverfahren die schulischen Kompetenzen beider Schülergruppen gleichermaßen reliabel und valide messen.

Neben der internen Struktur des Tests sollten die Zusammenhangsmuster der Kompetenztestwerte mit relevanten anderen Konstrukten für die betrachteten Schülergruppen ähnlich ausfallen. Die theoriegeleitete Überprüfung von Zusammenhängen der Testergebnisse mit externen Kriterien, wie z.B. den Ergebnissen anderer Leistungstests, ist ein wichtiger Aspekt der Validierung eines Testverfahrens (Messick, 1995). Zeigen sich ähnliche Zusammenhänge für die untersuchten Schülergruppen, können deren Leistungen auf einer gemeinsamen Skala abgebildet und sinnvoll vergleichend interpretiert werden.

4.1.2.1 Methodische Herausforderungen bei der Interpretation der Testwerte von Schülerinnen und Schülern mit SPF in LSAs

Die Übertragbarkeit der Interpretation von Testwerten auf Schülerinnen und Schüler mit SPF kann durch mehrere Aspekte eingeschränkt sein. Ein Aspekt ist die Reliabilität der Leistungstests, die bei dieser Schülergruppe möglicherweise geringer ist. Allgemein fällt die Messgenauigkeit von Leistungstests, die auf Modellen der *Item Response Theory* basieren, im mittleren Bereich der Fähigkeitsverteilung am höchsten aus und nimmt zu den Rändern der Fähigkeitsverteilung hin ab (Embretson & Reise, 2000). Für Schülerinnen und Schüler mit SPF ist im Vergleich zu Schülerinnen und Schülern ohne SPF eine deutlich geringere mittlere Fähigkeit (Abedi et al., 2011) sowie eine eingeschränkte Varianz in der Testleistung (Lane & Leventhal, 2015) zu erwarten. Dies kann die Messgenauigkeit der Leistungstests für Schülerinnen und Schüler mit SPF reduzieren.

Als weiterer Aspekt sind mögliche Einschränkungen der Testfairness in Betracht zu ziehen, die sich auf die Interpretierbarkeit der Testleistungen von Schülerinnen und Schülern mit SPF auswirken können. Die eingesetzten Testverfahren sollten es allen Schülerinnen und Schülern ermöglichen, ihre schulischen Kompetenzen zu zeigen. Treten unter Berücksichtigung von Fähigkeitsunterschieden noch differenzielle Schwierigkeiten der Items zwischen Schülergruppen auf, kann die vergleichende Interpretierbarkeit der Testwerte eingeschränkt sein (Abedi et al., 2011). Analysen zur sogenannten „differenziellen Itemfunktionalität“ (Differential Item Functioning, DIF, Holland & Thayer, 1988) können Hinweise darauf geben, inwieweit die Lösungswahrscheinlichkeit der einzelnen Items nicht nur durch die schulischen Kompetenzen beeinflusst wird, die erfasst werden sollen, sondern auch durch konstrukt-ferne Merkmale. DIF liegt vor, wenn Schülerinnen und Schüler aus verschiedenen Gruppen trotz gleichem Fähigkeitsniveau in der betrachteten Kompetenz ein spezifisches Item mit unterschiedlich hoher Wahrscheinlichkeit lösen. Bei den konstrukt-fernen Merkmalen, die zu systematischen Unterschieden in der Testbearbeitung und somit zu DIF führen können, kann es sich z.B. um sprachliche Anforderungen, Kontextualisierungen oder Antwortformate der Aufgaben handeln (Abedi et al., 2011; Heydrich et al., 2013).

4.1.2.2 Anpassung von Testverfahren für Schülerinnen und Schüler mit SPF in LSAs

Um die Eignung von in LSAs eingesetzten Leistungstests für Schülerinnen und Schülern mit SPF zu verbessern und den Einfluss konstrukt-irrelevanter Merkmale auf die Leistung bei einzelnen Items bzw. im Test insgesamt zu verringern, können sogenannte Akkommodationen vorgenommen werden (Pitoniak & Royer, 2001). Akkommodationen sind Anpassungen in der Testgestaltung, die dazu dienen sollen, diejenige konstrukt-irrelevante Varianz in den Testwerten zu reduzieren, die mit den Beeinträchtigungen der Kinder mit SPF zusammenhängt, ohne die Definition und Operationalisierung des zugrundeliegenden Konstrukts zu verändern (Elliott, Beddow, Kurz & Kettler, 2011).

Eine häufig eingesetzte Akkommodation betrifft die Testlänge. Die Teilnahme an Leistungstests in LSAs erfordert ein ausreichendes Maß an Konzentration und Aufmerksamkeit über einen längeren Zeitraum. Durch kürzere Tests für Schülerinnen und Schüler mit SPF können mögliche Nachteile dieser Schülergruppe in der Aufmerksamkeit und Konzentrationsfähigkeit berücksichtigt werden (Heydrich et al., 2013). Als Indikator dafür, ob die Bearbeitungszeit für Kinder mit SPF angemessen ist, kann die Anzahl fehlender Werte herangezogen werden. Insbesondere die Analyse fehlender Werte, die am Ende des Tests auftreten (*Missing not Reached*), kann Hinweise darauf geben, ob Schülerinnen und Schüler mit SPF durch die Länge des Tests in besonderem Maße überfordert waren.

Eine weitere Akkommodation, die in LSAs zum Einsatz kommt, besteht darin, besonders schwierige Items bzw. Testblöcke aus den Testheften für Schülerinnen und Schülern mit SPF zu entfernen (Südkamp, Pohl, Hardt, Jordan & Duchhardt, 2015). Durch den Einsatz leichter Aufgaben kann eine höhere Passung der Itemschwierigkeiten und Personenfähigkeiten erreicht werden, wodurch sich die Reliabilität der Kompetenztests für Schülerinnen und Schüler mit SPF erhöht (Lane & Leventhal, 2015).

4.1.2.3 Forschungsstand zur Interpretation der Testwerte von Schülerinnen und Schülern mit SPF in LSAs

Im Rahmen nationaler Erhebungen wurde nur selten überprüft, ob die Abbildung der Testwerte auf einer gemeinsamen Skala gerechtfertigt ist, da die Testscores von Heranwachsenden mit SPF und Heranwachsenden ohne SPF in vergleichbarer Weise interpretiert werden können. Eine Ausnahme ist die im Rahmen des Nationalen Bildungspanels (NEPS; Südkamp,

Pohl, Hardt et al., 2015; Südkamp, Pohl & Weinert, 2015) durchgeführte Zusatzerhebung an Förderschulen.

Südkamp, Pohl und Weinert (2015) untersuchten, inwieweit Leseverständnistests in NEPS für Schülerinnen und Schüler an Förderschulen mit dem Förderschwerpunkt *Lernen* eine reliable und vergleichbare Messung ermöglichen. Als Vergleichsgruppe wurden Schülerinnen und Schüler an Hauptschulen betrachtet. Beide Schülergruppen besuchten die fünfte Klasse. Außerdem wurde geprüft, inwieweit Testakkommodationen zu einer verbesserten Messung der Schülergruppe mit SPF führen. Die Testgüte wurde anhand des Anteils fehlender Werte, des Itemfits, der Reliabilität und anhand von DIF-Analysen evaluiert. Im Standardtest ohne Akkommodationen zeigten sich höhere Anteile fehlender Werte für Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* als bei Schülerinnen und Schülern ohne SPF. Im Hinblick auf den Itemfit wies ein bedeutsamer Anteil der Items für Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* eine schlechte Passung auf. Außerdem war die Reliabilität des Leseverständnistests deutlich geringer als für Hauptschülerinnen und -schüler. Im Hinblick auf die Testfairness zeigte sich bei einem bedeutsamen Anteil der eingesetzten Aufgaben ein starker DIF. Insgesamt war die Vergleichbarkeit der Testwerte zwischen den betrachteten Schülergruppen somit stark eingeschränkt. Der Einsatz von Akkommodationen (reduzierte Testzeit bzw. einfachere Aufgaben) führte zu einer verbesserten Abbildung der schulischen Kompetenzen (s. auch Südkamp, Pohl, Hardt et al., 2015).

Die beschriebene Studie zur Frage der äquivalenten Interpretierbarkeit der Testwerte bezog sich auf Schülerinnen und Schüler mit SPF in Förderschulen. Im Rahmen von LSA werden jedoch auch die Leistungen von Schülerinnen und Schüler mit SPF in sonstigen allgemeinbildenden Schulen erfasst. Zudem besteht erhebliches wissenschaftliches und bildungspolitisches Interesse an der Frage, inwieweit die Leistungsentwicklung von Schülerinnen und Schülern mit SPF durch die Art der Beschulung beeinflusst wird (Kocaj, Kuhl, Kroth, Pant & Stanat, 2014). Daher ist es wichtig zu prüfen, ob die eingesetzten Testverfahren Fähigkeitsunterschiede bei Kindern mit SPF in beiden Beschulungsarten gleichermaßen valide und reliabel abbilden können. Außerdem untersuchte die Mehrzahl der Studien Schülerinnen und Schüler in der Sekundarstufe, wohingegen sich die vorliegende Studie auf Schülerinnen und Schüler am Ende der Primarstufe bezieht.

4.1.3 Fragestellungen

In der vorliegenden Studie soll geprüft werden, inwieweit die im Rahmen des nationalen Bildungsmonitoring in der Primarstufe eingesetzten Schulleistungstests in den Fächern Deutsch und Mathematik dazu geeignet sind, die Kompetenzen von Schülerinnen und Schülern mit SPF in Förder- und Grundschulen zu erfassen (vgl. IQB-Ländervergleich 2011, Stanat, Pant, Böhme & Richter, 2012). In der 2011 durchgeführten Studie unterschieden sich die Testbedingungen zwischen den beiden Schularten: In Förderschulen kamen Testhefte mit einer reduzierten Anzahl von Aufgabenblöcken und einer entsprechend verkürzten Bearbeitungszeit zum Einsatz. Zudem enthielten die Testhefte eine Auswahl leichterer Aufgaben (s. Abschnitt 4.2). Auf der Grundlage dieser Daten werden vier Forschungsfragen untersucht.

1. Anteil fehlender Werte: Ist die Bearbeitungszeit für Kinder mit SPF angemessen? Anhand eines Vergleichs des Anteils fehlender Werte wird zunächst untersucht, ob die für die Kompetenztests zur Verfügung stehende Bearbeitungszeit für die drei Schülergruppen gleichermaßen ausreichend ist. Es wird angenommen, dass der Anteil fehlender Werte für Kinder mit SPF deutlich höher ist als für Kinder ohne SPF. Aufgrund der Anpassungen der in Förderschulen eingesetzten Tests sollte jedoch der Anteil nicht bearbeiteter Aufgaben bei Kindern mit SPF in Förderschulen geringer sein als bei Kindern mit SPF in Grundschulen.

2. Passung des Raschmodells: Treffen die Annahmen eines eindimensionalen Konstrukts für Kinder mit SPF in vergleichbarer Weise zu wie für Kinder ohne SPF? Im Hinblick auf die strukturelle Validität der Kompetenztests wird für die einzelnen Kompetenzdomänen geprüft, inwiefern sich die Annahme eines eindimensionalen Konstrukts in allen drei Schülergruppen bestätigen lässt. Wir nehmen an, dass sich auch die schulischen Kompetenzen der Kinder mit SPF in Förder- und Grundschulen mit dem Raschmodell in vergleichbarer Weise abbilden lassen.

3. Differenzielle Itemfunktionalität: Sind die eingesetzten Aufgaben für Kinder mit SPF differenziell schwieriger als für Kinder ohne SPF? Um Aspekte der Testfairness zu untersuchen, wird in einem nächsten Schritt geprüft, ob einzelne Items in den drei Gruppen nach Berücksichtigung von Fähigkeitsunterschieden differenzielle Schwierigkeiten aufweisen. Wir gehen davon aus, dass die Aufgaben für Kinder mit SPF – insbesondere in Förderschulen – zwar deutlich schwieriger sind als für Kinder ohne SPF, dass nach Kontrolle von Fähigkeits-

unterschieden zwischen den Schülergruppen darüber hinaus aber keine bedeutsamen Differenzen in den Itemschwierigkeiten bestehen.

4. Vergleich der Korrelationsmuster zwischen den Schülergruppen: Lässt sich die Interpretation der Testwerte im Hinblick auf Zusammenhänge mit externen Kriterien bei Kindern ohne SPF auf die Schülergruppe mit einem SPF übertragen? Abschließend wird ein Aspekt der externen Validität der Testverfahren geprüft, indem korrelative Zusammenhänge der Testwerte untereinander und mit den kognitiven Grundfähigkeiten zwischen den drei Schülergruppen verglichen werden. Für diese Variablen wird eine empirisch bewährte Korrelationsstruktur angenommen, die sich zwischen den drei Schülergruppen nicht bedeutsam unterscheiden sollte. So sollten die Kompetenzen im Lesen und Zuhören höher miteinander korrelieren als mit den mathematischen Kompetenzen.

4.1.4 Methoden

4.1.4.1 Stichprobe

Die in den folgenden Analysen verwendeten Daten wurden im Rahmen der 2011 durchgeführten Ländervergleichsstudie des IQB erhoben. Diese Querschnittsstudie diente der Überprüfung des Erreichens der Bildungsstandards am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Die gezogene Stichprobe der Viertklässlerinnen und Viertklässler ($N=27081$) war bundesweit und für jedes einzelne Bundesland repräsentativ.

Insgesamt umfasste die Stichprobe auch 1195 Schülerinnen und Schüler mit einem diagnostizierten SPF im Bereich *Lernen, Sprache* oder *emotionale und soziale Entwicklung* (durchschnittliches Alter: 10.90 Jahre, $SD = 0.64$, 66.3% Jungen). Die Mehrheit dieser Gruppe ($n=762$; 63.8%) besuchte eine Grundschule, während 433 Kinder mit SPF in Förderschulen unterrichtet wurden (36.2%). In die hier berichteten Analysen wurden nur Schülerinnen und Schüler einbezogen, die an den Kompetenztestungen teilgenommen hatten. Im Fach Deutsch bearbeiteten $N = 1071$ und im Fach Mathematik $N = 1053$ Schülerinnen und Schülern mit SPF die Kompetenztests (s. für eine differenzierte Darstellung nach Schulart und Förderschwerpunkt Tab. 2). Die Schülergruppe ohne SPF umfasste 25706 Schülerinnen und Schüler (durchschnittliches Alter: 10.43 Jahre, $SD = 0.49$, 50% Jungen).

Tabelle 2. *Stichprobenumfänge in den drei untersuchten Schülergruppen*

Schülergruppe	Fach	$N_{\text{insgesamt}}$	$N_{\text{Grundschule}}$	$N_{\text{Förderschule}}$
Schülerinnen und Schüler ohne SPF	Deutsch	25283	25283	–
	Mathematik	25266	25266	–
Schülerinnen und Schüler mit SPF ¹ insgesamt	Deutsch	1071	658	413
	Mathematik	1053	659	394
Schülerinnen und Schüler mit FS ² <i>Lernen</i>	Deutsch	536	278	258
	Mathematik	523	282	241
Schülerinnen und Schüler mit FS <i>Sprache</i>	Deutsch	301	166	135
	Mathematik	301	167	134
Schülerinnen und Schüler mit FS <i>emotionale und soziale Entwicklung</i>	Deutsch	234	214	20
	Mathematik	229	210	19

¹ sonderpädagogischer Förderbedarf, ² Förderschwerpunkt

4.1.4.2 Instrumente

Schulische Kompetenzen. Die schulischen Kompetenzen der Kinder wurden mit standardisierten Leistungstests in den Fächern Deutsch und Mathematik erfasst, wobei ein Multi-Matrix-Design zur Anwendung kam, d.h. jeder Schüler bekam ein Testheft mit einer Teilmenge aller Aufgaben. Die Testhefte sind so erstellt, dass innerhalb einer Klasse das Gesamt der Testaufgaben durch eine hinreichend große Teilmenge von Schülern bearbeitet wird, um gesicherte Aussagen über Leistungen auf Individual- und Klassenebene zu ermöglichen. Im Fach Deutsch wurden Kompetenzen in den Bereichen Lesen und Zuhören getestet. Der Kompetenztest im Fach Mathematik erfasste fünf inhaltsbezogene Teilkompetenzen, die zu einem Globalwert mathematischer Kompetenz zusammengefasst wurden.

Bei der Testdurchführung in Förderschulen kamen verschiedene Akkommodationen in den Testbedingungen zum Einsatz. Die Testhefte unterschieden sich in der Bearbeitungszeit sowie in der Anzahl und dem Schwierigkeitsgrad der eingesetzten Aufgaben. Die Testzeit betrug in Grundschulen 80, in Förderschulen 40 Minuten pro Fach. Damit einhergehend bearbeiteten Schülerinnen und Schüler in Förderschulen eine entsprechend geringere Anzahl von Aufgaben (s. Tab. 4). Die durchschnittliche Zeit, die den Kindern pro Item zur Verfügung stand, war in Förder- und Grundschulen vergleichbar. Außerdem wurde in Förderschulen eine Auswahl der leichteren Aufgaben aus dem gemeinsamen Aufgabenpool verwendet. Es wurden also keine neuen Testaufgaben für Förderschulen mit SPF entwickelt, sondern Aufgaben eingesetzt, die in Voruntersuchungen mit Kindern ohne SPF hohe Lösungswahrscheinlichkeiten aufwiesen. Somit waren die Testbedingungen für Kinder mit SPF je nach

Beschulungsart unterschiedlich. Schülerinnen und Schüler mit SPF in Grundschulen erhielten dieselben Testhefte wie Kinder ohne SPF und wurden unter denselben Bedingungen getestet, erfuhren also keine Akkommodationen.

Kognitive Grundfähigkeiten. Die kognitiven Grundfähigkeiten wurden mit zwei Untertests des Kognitiven Fähigkeitstests (KFT 4-12+R, Heller & Perleth, 2000) erhoben. Dabei diente die Subskala *Wortschatz* (Cronbachs $\alpha = .72$ für Kinder ohne SPF, $\alpha = .79$ für Kinder mit SPF in Grundschulen, $\alpha = .78$ für Kinder mit SPF in Förderschulen) als Indikator für die verbalen kognitiven Fähigkeiten und die Subskala *Figurenanalogien* ($\alpha = .92$ für Kinder ohne SPF, $\alpha = .92$ für Kinder mit SPF in Grundschulen, $\alpha = .88$ für Kinder mit SPF in Förderschulen) als Indikator für die nonverbalen kognitiven Fähigkeiten der Schülerinnen und Schüler.

4.1.4.3 Datenanalyse

Im Folgenden werden die Analysemethoden für jede der vier Forschungsfragen erläutert.

1. *Anteil fehlender Werte.* Zur Analyse fehlender Werte in den Kompetenztests wurden drei Kategorien unterschieden: *Missing by Intention*, *Missing Invalid Response* und *Missing not Reached*. Wenn eine Schülerin bzw. ein Schüler bei einem Item keine Antwort gegeben hatte, wurde die Kategorie *Missing by Intention* vergeben. Lag eine Antwort vor, war aber nicht interpretierbar, wurde diese Antwort als *Missing Invalid Response* kodiert. Die Kategorie *Missing not Reached* wurde für fehlende Itemantworten am Ende jedes Testblocks vergeben, wenn die letzten Items des Blocks durchgängig nicht beantwortet wurden und daher vermutet werden konnte, dass das Kind keine Möglichkeit hatte, die letzten Items zu bearbeiten.

2. *Passung des Raschmodells.* Zur separaten Analyse der Itemantworten in den drei betrachteten Schülergruppen wurde ein eindimensionales Modell der *Item Response Theory* (IRT) (Raschmodell, vgl. Embretson & Reise, 2000) verwendet. Dabei wurden in einem ersten Schritt unter Verwendung der *Marginal Maximum Likelihood* Schätzmethode die Itemschwierigkeiten bestimmt. In einem zweiten Schritt wurde unter Fixierung der Itemparameter auf die im ersten Schritt geschätzten Werte die Verteilung der latenten Personenfähigkeiten geschätzt. Fehlende Werte (*Missing by Intention*) in den Kompetenztests wurden sowohl bei der Schätzung der Itemparameter als auch bei der Schätzung der Personenfähigkeit als falsche Antworten kodiert.

Die Testitems wurden für jede betrachtete Schülergruppe separat skaliert; anschließend wurde die Passung des Modells für jede Schülergruppe geprüft. Zur Beurteilung der Passung wurde der standardisierte *Weighted Mean Square* (WMNSQ, Wright & Masters, 1982) für jedes Item herangezogen. Dieser Kennwert beschreibt die Abweichung der beobachteten Lösungswahrscheinlichkeit eines Items von der durch das Modell implizierten Wahrscheinlichkeit für ein bestimmtes Fähigkeitsniveau. Ein „Infit“ von 1 indiziert eine ideale Passung. In Anlehnung an das Vorgehen in PISA (OECD, 2012) wird ein WMNSQ von unter 0.80 bzw. von über 1.20 als starker Misfit gewertet. Ein WMNSQ von über 1.15 indiziert einen leichten Misfit (vgl. Pohl & Carstensen, 2012). Mit dieser Analyse wurde geprüft, ob die modellimplizierten Annahmen für jede der drei Schülergruppen zutreffen.

3. *Differenzielle Itemfunktionalität.* Im dritten Schritt wurden DIF-Analysen (Holland & Thayer, 1988) durchgeführt. Dabei wurden die geschätzten Itemparameter für die drei Schülergruppen paarweise miteinander verglichen, wobei für mittlere Fähigkeitsunterschiede zwischen den jeweiligen Gruppen kontrolliert wurde. Es wurde geprüft, ob sich die Itemschwierigkeiten zwischen den betrachteten Schülergruppen unter Berücksichtigung von Fähigkeitsunterschieden bedeutsam voneinander unterscheiden. Das Ausmaß des DIF pro Item wurde wie folgt kategorisiert (vgl. Penfield & Camilli, 2007): geringer DIF wenn $|DIF| < 0.43$ oder nicht signifikant > 0 ; moderater DIF wenn $|DIF| \geq 0.43$ und $|DIF|$ signifikant > 0 , sowie entweder $|DIF| < 0.64$ oder $|DIF|$ nicht signifikant > 0.43 ; starker DIF wenn $|DIF| \geq 0.64$ und signifikant > 0.43 .

4. *Vergleich der Korrelationsmuster zwischen Schülergruppen.* Um Hinweise auf die externe Validität der Testverfahren zu gewinnen, wurden im vierten Schritt korrelative Zusammenhänge der Testwerte in den Bereichen Lesen, Zuhören und Mathematik untereinander sowie mit verbalen und nonverbalen kognitiven Grundfähigkeiten zwischen den Schülergruppen verglichen. Um manifeste Indikatoren für die latente Personenfähigkeit zu gewinnen, wurden für jede Schülerin und jeden Schüler jeweils 15 *Plausible Values* pro Kompetenzbereich gezogen (von Davier, Gonzales & Mislevy, 2009). Zur leichteren Interpretierbarkeit wurden die Kompetenzwerte der Schülerinnen und Schüler auf die in LSAs übliche Berichtsmetrik ($M = 500$, $SD = 100$) transformiert. Die Analysen wurden für jeden der 15 Plausible Values pro Kompetenzbereich separat durchgeführt und anschließend nach Rubin (1987) zusammengefasst.

4.1.5 Ergebnisse

4.1.5.1 Forschungsfrage 1 – Anteil fehlender Werte

In Tabelle 3 sind die durchschnittlichen Anteile fehlender Werte an allen vorgelegten Testitems für die drei Schülergruppen in den Kompetenzbereichen Lesen, Zuhören und Mathematik abgebildet. Insgesamt unterscheidet sich der Anteil fehlender Werte zwischen den Gruppen erheblich. Bei Kindern ohne SPF war in allen untersuchten Bereichen der Anteil fehlender Werte durchgängig am kleinsten; bei Kindern mit SPF in Grundschulen war der Anteil jeweils ungefähr doppelt so hoch (Tab. 3). Schülerinnen und Schüler in Förderschulen hatten nahezu doppelt so viele fehlende Werte wie Kinder mit SPF in Grundschulen (Tab. 3). Insbesondere die Anteile fehlender Werte, die durch Auslassen (*Missing by Intention*) und durch Nichterreichen des Endes eines Testblocks (*Missing not Reached*) zustande kamen, waren in Mathematik und im Lesen bei Kindern mit SPF in Förderschulen deutlich höher als in den beiden anderen Schülergruppen (Tab. 3). Somit schienen insbesondere Schülerinnen und Schüler mit SPF in Förderschulen durch die Länge der Kompetenztests im Lesen und in Mathematik überfordert zu sein.

Tabelle 3. *Relativer Anteil fehlender Werte pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen*

	<i>Missing by Intention</i>	<i>Missing Invalid Response</i>	<i>Missing not Reached</i>	<i>Missings insgesamt</i>
Deutsch Lesen				
Kinder ohne SPF ¹	4.25%	0.79%	0.98%	6.02%
Kinder mit SPF in Grundschulen	9.05%	1.49%	2.28%	12.82%
Kinder mit SPF in Förderschulen	15.15%	2.32%	10.86%	28.33%
Deutsch Zuhören				
Kinder ohne SPF ¹	4.99%	0.99%	0.66%	6.65%
Kinder mit SPF in Grundschulen	8.47%	1.77%	1.97%	12.21%
Kinder mit SPF in Förderschulen	15.39%	2.19%	2.64%	20.22%
Mathematik				
Kinder ohne SPF ¹	6.85%	1.09%	2.06%	9.99%
Kinder mit SPF in Grundschulen	11.84%	1.71%	4.58%	18.14%
Kinder mit SPF in Förderschulen	19.17%	1.74%	12.27%	33.18%

¹ sonderpädagogischer Förderbedarf

4.1.5.2 Forschungsfrage 2 – Passung des Raschmodells

Die Ergebnisse der Analyse zur Passung des Raschmodells sind in Tabelle 4 dargestellt. Für Kinder ohne SPF ergab sich für nahezu alle Items eine zufriedenstellende Passung (Tab. 4). Lediglich im Kompetenzbereich Mathematik wiesen einige Items (2%) einen leichten Misfit auf. In der Gruppe der Kinder mit SPF in Grundschulen war der Anteil der Items mit einem Misfit vor allem in Mathematik höher (insgesamt 14%). In der Schülergruppe mit SPF in Förderschulen schließlich waren anteilig noch einmal deutlich mehr Mathematikitems mit einem Misfit zu verzeichnen als in den anderen Schülergruppen (insgesamt 25%). In den Kompetenzbereichen Lesen und Zuhören waren die Anteile der Items mit Misfit bei Kindern mit SPF in Förder- und Grundschulen etwas höher als bei Kindern ohne SPF (Tab. 4). Im Vergleich zum Kompetenztest in Mathematik lagen diese Anteile jedoch in einem akzeptablen Bereich (Tab. 4).

Außerdem wurde geprüft, ob die eingesetzten Testverfahren in allen Schülergruppen vergleichbar reliabel messen. Dies scheint der Fall zu sein: Die *Expected a posteriori* - Reliabilitäten (EAP-Reliabilitäten) als Maß für die durch das Raschmodell erklärte Varianz der Personenfähigkeiten im Verhältnis zur Gesamtvarianz der Personenfähigkeiten unterschieden sich in keinem der drei Kompetenzbereiche bedeutsam zwischen den Schülergruppen (Tab. 4).

Tabelle 4. *Itemfit pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen*

	Kinder ohne SPF ¹	Kinder mit SPF in GS ²	Kinder mit SPF in FS ³
Deutsch Lesen			
<i>N</i> Items: WMNSQ ⁴ < 0.80	1	0	0
<i>N</i> Items: WMNSQ > 1.15	0	1	1
<i>N</i> Items: WMNSQ > 1.20	0	4	1
<i>N</i> Items insgesamt	80	80	26
EAP-Reliabilität	.71	.73	.75
Deutsch Zuhören			
<i>N</i> Items: WMNSQ ⁴ < 0.80	0	0	0
<i>N</i> Items: WMNSQ > 1.15	0	1	1
<i>N</i> Items: WMNSQ > 1.20	0	0	0
<i>N</i> Items insgesamt	51	51	21
EAP-Reliabilität	.65	.68	.70
Mathematik			
<i>N</i> Items: WMNSQ ⁴ < 0.80	0	10	4
<i>N</i> Items: WMNSQ > 1.15	6	9	3
<i>N</i> Items: WMNSQ > 1.20	0	19	4
<i>N</i> Items insgesamt	275	275	44
EAP-Reliabilität	.91	.92	.89

¹ sonderpädagogischer Förderbedarf, ² Grundschulen, ³ Förderschulen,

⁴ standardisierter Weighted Mean Square (Wright & Masters, 1982);

Kriterien zur Beurteilung des Itemfits: starker Misfit: WMNSQ < 0.80 oder WMNSQ > 1.20, leichter Misfit: WMNSQ > 1.15 und WMNSQ < 1.20 (OECD, 2012; Pohl & Carstensen, 2012)

4.1.5.3 Forschungsfrage 3 – Differenzielle Itemfunktionalität

Im Folgenden werden die Ergebnisse zur differenziellen Itemfunktionalität getrennt für die paarweisen Gruppenvergleiche zwischen Kindern mit SPF in Grundschulen, Kindern mit SPF in Förderschulen und Kindern ohne SPF dargestellt (Tab. 5).

Tabelle 5. *Differenzielle Itemfunktionen pro Fach und Kompetenzbereich in den drei untersuchten Schülergruppen*

	Kinder mit SPF ¹ in GS ² vs. Kinder ohne SPF	Kinder in FS ³ vs. Kinder ohne SPF	Kinder mit SPF in FS vs. Kinder mit SPF in GS
Deutsch Lesen			
Linking-Konstante (Logit)	1.03	2.14	1.14
Linking-Konstante (BISTA ⁴)	95.84	199.61	106.31
<i>N</i> Items mit sign. DIF ⁵	19 (23.75%)	8 (30.77%)	5 (19.23%)
<i>N</i> Items mit geringem DIF	62 (77.5%)	20 (76.92%)	19 (73.08%)
<i>N</i> Items mit moderatem DIF	18 (22.5%)	5 (19.23)	7 (26.92%)
<i>N</i> Items mit starkem DIF	0 (0%)	1 (3.85%)	0 (0%)
<i>N</i> Items insgesamt	80	26	26
Deutsch Zuhören			
Linking-Konstante (Logit)	0.76	1.92	1.10
Linking-Konstante (BISTA ⁴)	84.19	213.63	122.51
<i>N</i> Items mit sign. DIF ⁵	16 (31.37%)	8 (38.10%)	5 (23.81%)
<i>N</i> Items mit geringem DIF	41 (80.39%)	14 (66.67%)	13 (61.90%)
<i>N</i> Items mit moderatem DIF	10 (19.61%)	6 (28.57%)	8 (38.10%)
<i>N</i> Items mit starkem DIF	0 (0%)	1 (4.76%)	0 (0%)
<i>N</i> Items insgesamt	51	21	21
Mathematik			
Linking-Konstante (Logit)	1.05	2.19	1.08
Linking-Konstante (BISTA ⁴)	96.11	199.77	98.78
<i>N</i> Items mit sign. DIF ⁵	60 (21.82%)	26 (59.09%)	10 (22.73%)
<i>N</i> Items mit geringem DIF	220 (80.00%)	25 (56.82%)	33 (75%)
<i>N</i> Items mit moderatem DIF	54 (19.63%)	17 (38.64%)	10 (22.73%)
<i>N</i> Items mit starkem DIF	1 (0.36%)	2 (4.54%)	1 (2.27%)
<i>N</i> Items insgesamt	275	44	44

¹sonderpädagogischer Förderbedarf, ²Grundschulen, ³Förderschulen, ⁴Bildungsstandards-Metrik,

⁵Differential Item Functioning mit folgenden Kriterien zur Beurteilung des Ausmaß an DIF: geringer DIF: $|DIF| < 0.43$ oder nicht signifikant > 0 , moderater DIF: $0.43 < |DIF| < 0.64$ und $|DIF|$ signifikant > 0 , starker DIF: $|DIF| > 0.64$ und signifikant > 0.43 (Penfield & Camilli, 2007)

Kinder mit SPF in Grundschulen vs. Kinder ohne SPF

Schülerinnen und Schüler mit SPF in Grundschulen erhielten dieselben Testhefte wie Schülerinnen und Schüler ohne SPF. Daher konnte in allen Kompetenzbereichen der gesamte Itempool auf DIF überprüft werden. Die Ergebnisse sind in Tabelle 5 zusammengefasst. Für alle drei Kompetenzbereiche zeigte sich erwartungsgemäß, dass die Items von Kindern mit SPF deutlich weniger häufig gelöst wurden als von Kindern ohne SPF (s. Linking-Konstante in Tab. 5). Im Lesetest wies keines der 80 Items einen starken DIF auf; bei 18 Items (23%) wurde ein moderater DIF festgestellt. Im Kompetenzbereich Zuhören wies keines der 51 Items einen starken DIF auf; bei 10 Items (20%) lag ein moderater DIF vor. Ein ähnliches Muster ergab sich für den Kompetenzbereich Mathematik, in dem nur 1 Item mit starkem und 54 Items (20%) mit moderatem DIF zu verzeichnen waren.

Kinder mit SPF in Förderschulen vs. Kinder ohne SPF

Im Gegensatz zu Kindern mit SPF in Grundschulen erhielten Schülerinnen und Schüler in Förderschulen verkürzte Testhefte. Eine vergleichende Analyse der Itemschwierigkeiten kann somit nur für diejenigen Items vorgenommen werden, die beiden Schülergruppen vorgelegt wurden. Im Vergleich zu Kindern ohne SPF waren die Items für Kinder in Förderschulen über alle Kompetenzbereiche deutlich schwieriger (Tab. 5). Der Anteil an Items mit einem starken DIF war zwar höher als im Vergleich zwischen Kindern mit SPF in Grundschulen und Kindern ohne SPF, die Anzahl der Testaufgaben, die dies betraf, war aber ebenfalls gering.

Kinder mit SPF in Förderschulen vs. Kinder mit SPF in Grundschulen

Auch bei den DIF-Analysen zum Vergleich von Kindern mit SPF in Förder- vs. Grundschulen konnte nur auf die Teilmenge des Itempools zurückgegriffen werden, die beide Schülergruppen bearbeitet hatten. Für Kinder mit SPF in Grundschulen waren die Aufgaben in allen Kompetenzbereichen leichter zu lösen als für Kinder in Förderschulen (Tab. 5). Die Differenzen in den Itemschwierigkeiten waren im Vergleich dieser Gruppen ähnlich groß wie die Unterschiede zwischen Kindern ohne SPF und Kindern mit SPF in Grundschulen (Tab. 5). Über alle Kompetenzbereiche hinweg wies jedoch lediglich ein Mathematikitem einen starken DIF auf. Der Anteil an Items mit moderatem DIF war dagegen etwas höher. Im Lesen zeigten 7 von 26 Items (27%), im Zuhören 8 von 21 Items (38%) und in Mathematik 10 von 44 Items (23%) einen moderaten DIF. Der geringe Anteil von Items mit einem starken DIF weist darauf hin, dass in den drei Kompetenzbereichen für Kinder mit SPF in Förder- und Grund-

schulen nach Kontrolle der Personenfähigkeiten keine substanziellen Unterschiede in den Itemschwierigkeiten auftraten, die Items demnach keine Gruppe systematisch „benachteiligten“.

4.1.5.4 Forschungsfrage 4 – Vergleich der Korrelationsmuster zwischen den Schülergruppen

Im Hinblick auf die externe Validität der Testverfahren wurde geprüft, ob die Zusammenhänge der Leistungen im Lesen, Zuhören und in Mathematik untereinander sowie zu den verbalen und nonverbalen kognitiven Grundfähigkeiten zwischen den Schülergruppen ähnlich ausgeprägt sind. Tabelle 6 stellt die Ergebnisse der Korrelationsanalysen dar. In allen drei Gruppen korrelierten die Testwerte im Lesen, Zuhören und in Mathematik stark positiv miteinander. Allerdings waren die Korrelationen zwischen den sprachlichen Kompetenzbereichen und Mathematik bei Kindern in Förderschulen höher als in den anderen Schülergruppen (Tab. 6). Die Testwerte im Lesen und Zuhören hingen in den drei Schülergruppen stärker mit den verbalen kognitiven Grundfähigkeiten zusammen als mit den nonverbalen kognitiven Grundfähigkeiten (Tab. 6). Auch in Mathematik korrelierten die Testwerte stärker mit den verbalen als mit den nonverbalen kognitiven Grundfähigkeiten, jedoch waren die Zusammenhänge mit den nonverbalen kognitiven Grundfähigkeiten in Mathematik enger als mit den beiden Kompetenzbereichen im Fach Deutsch.

Mit paarweisen Vergleichen wurde geprüft, ob sich die Korrelationen zwischen den Gruppen signifikant voneinander unterscheiden (Tab. 6). Bei 6 von 27 Vergleichen fanden sich statistisch bedeutsame Unterschiede auf einem Signifikanzniveau von $\alpha = .05$. Diese betreffen in allen Fällen die Schülergruppe mit SPF in Förderschulen: Bei 4 Vergleichen zeigten sich höhere Korrelationen für Kinder mit SPF in Förderschulen als für Kinder ohne SPF; bei 2 Vergleichen waren die Korrelationen für diese Schülergruppe höher als für Kinder mit SPF in Grundschulen (Tab. 6).

Tabelle 6. Korrelationen zwischen schulischen Kompetenzen und kognitiven Grundfähigkeiten in den drei untersuchten Schülergruppen

	1. Lesen ¹			2. Zuhören			3. Mathematik		
	Kinder ohne SPF ²	Kinder mit SPF in GS ³	Kinder mit SPF in FS ⁴	Kinder ohne SPF	Kinder mit SPF in GS	Kinder mit SPF in FS	Kinder ohne SPF	Kinder mit SPF in GS	Kinder mit SPF in FS
1. Lesen									
2. Zuhören	.72	.71	.69						
3. Mathematik	.66 ^a	.65 ^b	.72 ^{a, b}	.62 ^a	.61 ^b	.68 ^{a, b}			
4. KFT ⁵ Verbal	.68	.71	.68	.68 ^a	.70	.74 ^a	.61 ^a	.63	.69 ^a
5. KFT Nonverbal	.46	.41	.41	.45	.41	.47	.53	.49	.54

¹ Testleistungen (erfasst über je 15 Plausible Values), ² sonderpädagogischer Förderbedarf,

³ Grundschulen, ⁴ Förderschulen, ⁵ Kognitiver Fähigkeitstest

^a signifikanter Unterschied ($p < .05$, zweiseitige Testung) in der Korrelation zwischen Kindern ohne SPF und Kindern mit SPF in Förderschulen

^b signifikanter Unterschied ($p < .05$, zweiseitige Testung) in der Korrelation zwischen Kindern mit SPF in Förderschulen und Kindern mit SPF in Grundschulen

4.1.6 Diskussion

4.1.6.1 Zusammenfassung und Interpretation

Die Ergebnisse der vorliegenden Analysen weisen darauf hin, dass eine reliable und valide Testung von Schülerinnen und Schülern mit SPF in LSAs im Prinzip möglich ist. Einzelne Unterschiede in der Testgüte waren in der untersuchten Studie in erster Linie zwischen Kindern an Förderschulen und Kindern ohne SPF an Grundschulen zu verzeichnen.

Kinder mit SPF in Grundschulen

Für die Schülergruppe mit SPF in Grundschulen deuten die vorliegenden Ergebnisse darauf hin, dass sich die schulischen Kompetenzen auf einer gemeinsamen Skala mit den Testwerten der Schülergruppe ohne SPF abbilden lassen. Diese Kinder erhielten dieselben Aufgaben wie Schülerinnen und Schüler ohne SPF und bearbeiteten die Kompetenztests ohne Akkommodationen. Bei Anwendung des Raschmodells konnte eine akzeptable Passung der eingesetzten Aufgaben für Schülerinnen und Schüler mit SPF an Grundschulen in allen drei Kompetenzbereichen festgestellt werden.

Allerdings war der Anteil fehlender Werte in den drei Kompetenzbereichen deutlich höher als für Kinder ohne SPF. Diese Ergebnisse weisen darauf hin, dass die eingesetzten Tests für Schülerinnen und Schüler mit SPF zu umfangreich sind, um sie in der vorgegebenen Zeit vollständig beantworten zu können. Die Vergleichbarkeit der Testwerte durch differenzielle Anteile fehlender Werte wäre vor allem dann eingeschränkt, wenn diese mit konstrukt-irrelevanten Merkmalen zusammenhängen.

Die Ergebnisse der DIF-Analysen zeigen, dass Kinder mit SPF in Grundschulen durch die eingesetzten Aufgaben kaum eine Messbenachteiligung erfahren haben. Zudem ergaben Korrelationsanalysen für diese Gruppe ähnliche Zusammenhangsmuster der Testwerte im Leseverständnis, Zuhören und in Mathematik wie für Schülerinnen und Schüler ohne SPF. Ebenso korrelierten die kognitiven Grundfähigkeiten in vergleichbarer Stärke mit den Leistungen in den Kompetenztests. Diese Ergebnisse sprechen dafür, dass Testwerte von Schülerinnen und Schülern mit SPF in Grundschulen und Kindern ohne SPF in gleicher Weise das Kompetenzniveau reflektieren, das den Tests zugrunde liegt.

Kinder mit SPF in Förderschulen

Für die Schülerinnen und Schüler mit SPF in Förderschulen, die die Testverfahren mit den eingesetzten Akkommodationen bearbeiteten, scheint die Vergleichbarkeit der Messungen deutlich stärker eingeschränkt zu sein. Für die Kompetenzbereiche Lesen und Zuhören lässt sich anhand der Ergebnisse aus den Itemfit-Analysen schlussfolgern, dass das Raschmodell geeignet ist, auch die Antworten der Kinder mit SPF in Förderschulen auf die Items angemessen abzubilden. Im Kompetenzbereich Mathematik wies hingegen ein beträchtlicher Anteil der Items in dieser Gruppe einen Misfit auf. Die Abbildung der mathematischen Kompetenzen der Schülerinnen und Schüler an Förderschulen auf einer gemeinsamen Skala mit Schülerinnen und Schülern ohne SPF ist daher nur eingeschränkt möglich.

Erwartungsgemäß zeigte sich in allen betrachteten Kompetenzbereichen, dass die Aufgaben für Schülerinnen und Schüler mit SPF in Förderschulen deutlich schwieriger waren als für Kinder ohne SPF. Die Schwierigkeitsunterschiede zwischen Schülerinnen und Schülern mit SPF in Förder- und Grundschulen waren dagegen geringer ausgeprägt. Werden Fähigkeitsunterschiede zwischen den Schülergruppen berücksichtigt, legen die Ergebnisse der DIF-Analysen nahe, dass die Gütekriterien für die Population der Kinder mit SPF in Förderschulen durch die eingesetzten Aufgaben nicht systematisch eingeschränkt sind. Insbesondere im Vergleich der schulischen Kompetenzen von Kindern mit SPF in Förder- und Grundschulen

war der Anteil an Items mit einem starken DIF vernachlässigbar. Im Hinblick auf Aspekte der externen Validität der eingesetzten Kompetenztests zeigten sich für Kinder mit SPF in Förderschulen leicht abweichende Zusammenhänge der Testwerte im Leseverständnis, Zuhören und Mathematik im Vergleich zu Schülerinnen und Schülern mit SPF in Grundschulen bzw. der Schülergruppe ohne SPF. Die Testwerte in Mathematik korrelierten in dieser Schülergruppe stärker mit den Kompetenzen im Lesen und Zuhören als in den anderen Gruppen. Diese Ergebnisse schränken die Interpretation der Testwerte von Kindern in Förderschulen im Vergleich zu Schülerinnen und Schülern mit SPF in Grundschulen und Kindern ohne SPF möglicherweise ein.

4.1.6.2 Ausblick

Die von Schülerinnen und Schülern mit SPF in LSAs erreichten Leistungen werden durch deren Kompetenzen und individuelle Lernvoraussetzungen, aber auch durch Merkmale der Testaufgaben beeinflusst (Lane & Leventhal, 2015). Dieses Zusammenspiel zwischen Schüler- und Testmerkmalen sollte in zukünftigen Studien stärker berücksichtigt werden. Zum einen ist die Heterogenität der Schülerschaft mit SPF auch innerhalb der Schularten in Betracht zu ziehen. In den vorliegenden Analysen konnte beispielsweise keine differenzierte Betrachtung nach den verschiedenen Förderschwerpunkten vorgenommen werden, da die Teilstichproben zu klein waren. Eine entsprechende Analyse könnte Hinweise darauf geben, ob für Schülerinnen und Schüler in Abhängigkeit vom Förderschwerpunkt unterschiedliche Akkommodationen bereitgestellt werden sollten (Elliott et al., 2011). Außerdem dürften zwischen den Förderschwerpunkten Unterschiede dahingehend bestehen, ob die Schülerinnen und Schüler zielgleich oder zieldifferent unterrichtet werden (Autorengruppe Bildungsberichterstattung, 2014). Zieldifferent unterrichtete Schülerinnen und Schüler mit SPF haben spezielle Lehrpläne und werden daher mit den Anforderungen der an den Bildungsstandards ausgerichteten Kompetenztests weniger vertraut sein. Dies könnte die Vergleichbarkeit der Testwerte zwischen zielgleich und zieldifferent unterrichteten Schülerinnen und Schülern einschränken (Elliott et al., 2011).

Zum anderen sollte in weiterführenden Analysen geprüft werden, welche Aufgabenmerkmale (z.B. sprachliche Komplexität, Antwortformat) mit differenzieller Itemfunktionalität für Heranwachsende mit SPF zusammenhängen (vgl. Haag, Heppt, Stanat, Kuhl & Pant, 2013).

Damit könnten potenzielle konstrukt-fremde Merkmale der Testaufgaben identifiziert werden, die bei der Aufgabenentwicklung berücksichtigt werden sollten.

Die vorliegende Untersuchung deutet darauf hin, dass der Einbezug von Schülerinnen und Schülern mit SPF in LSAs im Bildungsbereich prinzipiell möglich, teilweise jedoch mit Herausforderungen für die Testgestaltung verbunden ist. Die schulischen Kompetenztests müssen auf der einen Seite die Fähigkeiten der gesamten Schülerpopulation reliabel und vergleichbar abbilden. Auf der anderen Seite sollten die Leistungsstände auch von Schülerinnen und Schülern mit SPF so akkurat und differenziert gemessen werden, dass die Ergebnisse auch für die Evaluation von Inklusionsbemühungen herangezogen werden können.

Die Generalisierbarkeit der Interpretation der Testwerte wird in der vorliegenden Untersuchung dadurch eingeschränkt, dass die eingesetzten Aufgaben für Kinder mit SPF im Durchschnitt deutlich schwieriger waren als für Kinder ohne SPF. Dadurch können Fähigkeitsunterschiede innerhalb der Schülergruppe mit SPF in Förderschulen im unteren Fähigkeitsbereich nicht mehr ausreichend differenziert werden (Südkamp, Pohl, Hardt, et al., 2015). Die vorliegenden Ergebnisse zeigen, dass Testverfahren, die nicht gezielt für diese Schülergruppe entwickelt und erprobt wurden, insbesondere im Fach Mathematik in Förderschulen zu schwer sind, um auch innerhalb dieser Gruppe eine reliable Interpretation der Testwerte zu ermöglichen. Dieser Herausforderung könnte durch die gezielte Entwicklung und Erprobung besonders leichter, auf die Gruppe der Schülerinnen und Schüler mit SPF abgestimmter Testaufgaben begegnet werden.

Literaturverzeichnis

- Abedi, J., Leon, S., Kao, J., Bayley, R., Ewers, N., Herman, J., & Mundhenk, K. (2011). *Accessible reading assessments for students with disabilities: The role of cognitive, grammatical, lexical, and textual/visual features. CRESST Report 785*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Böhme, K., & Robitzsch, A. (2009). Methodische Aspekte der Erfassung der Lesekompetenz. In D. Gränzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss, & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 250–289). Weinheim: Beltz.
- Bos, W., Müller, S., & Stubbe, T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Chudowsky, N., & Pellegrino, J. W. (2003). Large-scale assessments that support learning: What will it take? *Theory Into Practice*, 42, 75–83.
- Cole, N. S., & Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38, 369–382.
- Elliott, S. N., Beddow, P. A., Kurz, A., & Kettler, R. J. (2011). Creating access to instruction and tests of achievement: Challenges and solutions. In S. N. Elliott, R. J. Kettler, P. A. Beddow, & A. Kurz (Hrsg.), *Handbook of accessible achievement tests for all students* (S. 1–16). New York, NY: Springer.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Erlbaum: Hillsdale, NJ.
- Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, 28, 24–34.
- Heller, K. A., & Perleth, C. (2000). *KFT 4-12+R: kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test.
- Heydrich, J., Weinert, S., Nusser, L., Artelt, C., & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: Challenges and approaches within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online*, 5, 217–240.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Hrsg.), *Test validity* (S. 129–145). Hillsdale, NJ: Erlbaum.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A., & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191.
- Kornmann, R. (2006). Die Überrepräsentation ausländischer Kinder und Jugendlicher in Sonderschulen mit dem Schwerpunkt Lernen. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (S. 71–85). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lane, S., & Leventhal, B. (2015). Psychometric challenges in assessing English language learners and students with disabilities. *Review of Research in Education*, 39, 165–214.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5–8.
- OECD. (2012). *PISA 2009 technical report*. Paris: OECD Publishing.
- Olson, J. F., & Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress* (No. NCES 97-482). U.S. Department of Education, Office of Educational Research and Improvement.
- Opp, G., Budnik, I., & Fingerle, M. (2008). Sonderschulen — integrative Beschulung. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 341–361). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Penfield, R. D., & Camilli, G. (2007). Differential item functioning and item bias. In C.R. Rao and S. Sinharay (Hrsg.), *Handbook of Statistics* (Bd. 26, S. 125–167). Amsterdam: Elsevier/North Holland.
- Pitoniak, M. J., & Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, 71, 53–104.
- Pohl, S., & Carstensen, C. H. (2012). *NEPS technical report – Scaling the data of the competence tests*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Ricking, H. (2005). Der „Overlap“ von Lern- und Verhaltensstörungen. *Sonderpädagogik*, 35, 235–248.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons.
- Smith, E., & Douglas, G. (2014). Special educational needs, disability and school accountability: an international perspective. *International Journal of Inclusive Education*, 18, 443–458.

- Stanat, P., Pant, H. A., Böhme, K., & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Südkamp, A., Pohl, S., Hardt, K., Jordan, A.-K., & Duchhardt, C. (2015). Kompetenzmessung in den Bereichen Lesen und Mathematik bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 243–272). Wiesbaden: Springer VS.
- Südkamp, A., Pohl, S., & Weinert, S. (2015). Competence assessment of students with special educational needs - Identification of appropriate testing accommodations. *Frontline Learning Research*, 3(2), 1–25.
- Thurlow, M. L. (2002). Positive educational results for all students: The promise of standards-based reform. *Remedial and Special Education*, 23, 195–202.
- Thurlow, M. L. (2010). Steps toward creating fully accessible reading assessments. *Applied Measurement in Education*, 23, 121–131.
- UN-BRK. Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention) vom 21. 12. 2008.
- Von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series Volume*, 2, 9–36.
- Winkelmann, H., & Robitzsch, A. (2009). Modelle mathematischer Kompetenzen: Empirische Befunde zur Dimensionalität. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss, & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 169–196). Weinheim: Beltz.
- Wrase, M. (2015). Die Implementation des Rechts auf inklusive Schulbildung nach der UN-Behindertenrechtskonvention und ihre Evaluation aus rechtlicher Perspektive. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 41–74). Wiesbaden: Springer VS.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis. Rasch measurement*. Chicago, IL: MESA Press.

4.2 Teilstudie 2: Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe

Abstract

Im Zuge der Umsetzung der UN-Behindertenrechtskonvention werden zunehmend Bemühungen unternommen, Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) inklusiv bzw. integrativ an Regelschulen zu unterrichten. Die vorliegende Studie untersucht im Primarbereich, ob sich die Kompetenzen von Schülerinnen und Schülern mit SPF an Regel- und Förderschulen unterscheiden. Anhand von Daten des bundesweiten *IQB-Ländervergleichs Primarstufe 2011* werden die mit standardisierten Leistungstests erhobenen Kompetenzen von Kindern mit SPF an Regelschulen ($n = 658$) und an Förderschulen ($n = 413$) verglichen. Um für bestehende Gruppenunterschiede in Merkmalen zu kontrollieren, die für Schulleistungen relevant sind, wird ein *Propensity Score Matching*-Verfahren eingesetzt. Nach dem Matching zeigt sich, dass Kinder mit SPF an Grundschulen signifikant höhere Kompetenzwerte im Lesen, Zuhören und in Mathematik aufweisen als vergleichbare Schülerinnen und Schüler an Förderschulen. Die Effekte sind besonders ausgeprägt für Kinder mit dem Förderschwerpunkt *Lernen*, weniger prägnant hingegen für Kinder mit dem Förderschwerpunkt *Sprache*. Mögliche Erklärungsansätze für die höheren schulischen Kompetenzen von Kindern mit SPF in Regelschulen und Unterschiede in Abhängigkeit vom Förderschwerpunkt werden diskutiert.

Schlagwörter: Inklusion, sonderpädagogischer Förderbedarf, Primarstufe, schulische Kompetenzen, Propensity Score Matching

Teilstudie 2 ist in folgender Publikation erschienen:

Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191. doi: 10.1007/s11577-014-0253-x

4.2.1 Einleitung

In Deutschland wird nicht erst seit der Verabschiedung der UN-Behindertenrechtskonvention (BRK 2008) gefordert, allen Schülerinnen und Schülern Zugang zu den gleichen Bildungseinrichtungen zu ermöglichen. Schon zu Beginn der 1990er Jahre wurden erste Versuche unternommen, die gängige Praxis separater Beschulung von behinderten Kindern in Förderschulen aufzugeben und diese zunehmend an Regelschulen zu unterrichten (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, KMK 1994). Dieser Wandel spiegelt sich auch in der begrifflichen Umorientierung von der *Sonderschulbedürftigkeit* hin zum *sonderpädagogischen Förderbedarf* (SPF) wider (KMK 1994; Preuss-Lausitz 2001). Während die ursprüngliche Klassifizierung von Kindern als sonderschulbedürftig eng mit der Beschulung an Sonder- bzw. Förderschulen verknüpft war, ist mit dem Konzept des sonderpädagogischen Förderbedarfs eine individualisierte Ausrichtung im Umgang mit behinderten Schülerinnen und Schülern verbunden. In Anlehnung an den Artikel 24 der BRK empfiehlt die KMK (2010) den Ausbau des gemeinsamen Lernens von Kindern und Jugendlichen mit und ohne Behinderung in der Regelschule und definiert somit die gemeinsame Bildung und Erziehung Heranwachsender als zentrales Ziel bildungspolitischer Maßnahmen.

Im Zuge der Umstellung zu einer integrativen bzw. inklusiven Beschulung stellt sich die Frage, wie sich die Art der Beschulung (Regel- oder Förderschule) auf die schulischen Kompetenzen der Schülerinnen und Schüler mit SPF auswirkt.⁶ Befürworter einer separaten Beschulung argumentieren, dass Förderschulen in besonderem Maße geeignet seien, auf die speziellen Bedürfnisse von Kindern mit SPF einzugehen und ihnen damit optimale Bildungswege zu ermöglichen. Fürsprecher einer gemeinsamen Beschulung von Kindern mit SPF in Regelschulen hingegen führen an, dass diese den Schülerinnen und Schülern ein besonders anregendes Lernumfeld bieten und Bildungschancen eröffnen, die ihnen an Förderschulen

⁶ Integration und Inklusion stellen zwei unterschiedliche pädagogische Konzepte dar. Bei der Integration liegt der Fokus auf der Anpassung der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf an die Lernbedingungen in Regelschulen, z. B. durch die Schaffung spezieller Integrationsklassen (Hinz 2002). Im Gegensatz dazu fordern Vertreter der Inklusion eine Gestaltung der Lernumwelten in Regelschulen, die die Bedürfnisse von Kindern mit SPF und darüber hinaus die Heterogenität der gesamten Schülerschaft berücksichtigt (Biewer 2010). In der vorliegenden Untersuchung kann nicht zwischen beiden Konzepten differenziert werden.

zumeist verwehrt blieben. Der Zusammenhang zwischen Art der Beschulung und schulischen Kompetenzen von Kindern mit SPF wurde in Deutschland bislang jedoch nicht systematisch anhand von Daten großer nationaler Stichproben untersucht.

4.2.2 Theoretischer Hintergrund und Forschungsstand

Die Frage nach der optimalen Lernumgebung für Kinder mit SPF wird in der Forschung kontrovers diskutiert. In diesem Abschnitt sollen zum einen Argumente dafür beschrieben werden, dass Schülerinnen und Schüler mit SPF von einer gemeinsamen Beschulung in Regelschulen profitieren können. Zum anderen wird auf mögliche Gründe eingegangen, die für eine Beschulung von Kindern mit SPF in Förderschulen sprechen. Abschließend wird der internationale und nationale Forschungsstand zu Schulleistungen von Kindern mit SPF in Abhängigkeit von der Art der Beschulung dargestellt.

4.2.2.1 Theoretische Annahmen zur Auswirkung der Art der Beschulung auf die schulischen Kompetenzen von Kindern mit SPF

Die Einrichtung und der Ausbau der Sonder- bzw. Förderschulen in Deutschland waren geprägt von Bemühungen, die heterogenen Bedürfnisse von Kindern mit SPF angemessen zu berücksichtigen und ihnen eine optimale Schulbildung zu ermöglichen. Die Ausdifferenzierung verschiedener Schularten im deutschen Schulwesen basiert auf der Annahme, dass durch die Schaffung homogener Lerngruppen die Lernentwicklung von Heranwachsenden mit unterschiedlichen Lernvoraussetzungen optimiert werden kann. Solche Differenzierungsbemühungen können jedoch auch zu schulartspezifischen Lern- und Entwicklungsmilieus führen, die sich in Bezug auf institutionelle Faktoren und die Zusammensetzung der Schülerschaft unterscheiden und für die Lernentwicklung von Schülerinnen und Schülern relevant sind (Baumert et al. 2006; Stanat et al. 2012). Überträgt man diese theoretischen Annahmen auf den Kontext der vorliegenden Untersuchung, könnten Kinder mit SPF in Regelschulen eine im Vergleich zu Förderschulen anregendere Lernumgebung vorfinden, die ihre schulische Kompetenzentwicklung begünstigt (Bos et al. 2010; Wocken 2005). Diese Effekte könnten zum einen durch institutionelle Unterschiede zwischen Regel- und Förderschulen verursacht werden, die sich in schulartspezifischen Lehrplänen, didaktischen Traditionen, Aspekten der Lehrerbildung und der Unterrichtsorganisation widerspiegeln (Baumert et al. 2006; 2009; Hallinan & Kubitschek 1999; Hattie 2002). Verschiedene Autorinnen und

Autoren nehmen an, dass die Lehrpläne und die Unterrichtsgestaltung in Regelschulen im Vergleich zu Förderschulen stärker auf den akademischen Fortschritt der Klasse ausgerichtet sind (Hocutt 1996; Schumann 2007; Wocken 2000; 2005). In diesem Zusammenhang konnte festgestellt werden, dass der Einsatz herausfordernder Aufgaben im Unterricht schulische Leistungen fördert, während eine repetitive Unterrichtsführung und Reduzierung des Unterrichtstempos mit einer weniger günstigen Kompetenzentwicklung einhergeht (Baumert et al. 1986). Wocken (2005, S.62) verweist auf den geringen inhaltlichen Umfang der Lehrpläne für Förderschulen mit dem Förderschwerpunkt *Lernen* und schlussfolgert daraus, dass diese nicht dazu geeignet seien, Leistungs Nachteile von Kindern mit SPF zu kompensieren (s. auch Markussen 2004).

Zum anderen kann der Einfluss der Lernumgebung durch die leistungsbezogene und soziale Schülerzusammensetzung in Klassen (Maaz et al. 2008; Wocken 2000) und interaktionale Prozesse im Unterricht (Hocutt 1996; Schumann 2007; Wocken 2005) vermittelt werden. Dabei wird angenommen, dass vor allem leistungsschwächere Schülerinnen und Schüler vom Unterricht in leistungsheterogenen Lerngruppen profitieren (Lehmann 2006; Scharenberg 2012; Slavin 1996). So könnten in solchen Kontexten leistungsstärkere Mitschülerinnen und Mitschüler als positive Lernvorbilder fungieren und die Übernahme erfolgreicher Lernstrategien ermöglichen (Scharenberg 2012; Slavin 1990). Zudem könnte die potenziell höhere Leistungserwartung von Lehrkräften in Regelschulen die Kompetenzentwicklung von Kindern mit SPF fördern (Dar und Resh 1986; Gamoran 1986; Hornstra et al. 2010).

Darüber hinaus wird angenommen, dass mit der Überweisung von Kindern mit SPF auf Förderschulen eine Verstärkung sozialer Ungleichheiten im Bildungssystem einhergeht (Bos et al. 2010). So konnten empirische Studien zeigen, dass Kinder in Förderschulen (v.a. mit dem Förderschwerpunkt *Lernen*) eine sozial benachteiligte Gruppe darstellen (Dietze 2011; Wocken 2000, 2005). Die Überrepräsentation von Kindern aus sozial schwachen Familien an Förderschulen scheint dabei nicht nur auf primäre Herkunftseffekte (Boudon 1974) zurückzuführen sein, sondern auch auf sekundäre Herkunftseffekte, die u.a. durch subjektive Komponenten bei der Diagnostik eines SPF und der damit verbundenen Entscheidung über die Art der Beschulung entsteht (Gomolla 2006; Lentz & Radtke 1994). Die aus diesen Mechanismen resultierenden Unterschiede in der sozio-kulturellen Klassenkomposition zwischen Regel- und Förderschulen könnten die Qualität des Unterrichts und die Interaktion zwischen Lehrkräften und Schülerinnen und Schülern beeinflussen (Agirdag et al. 2012; Barth et al. 2004; Maaz et al. 2010). Dadurch könnten möglicherweise differenzielle Lern- und Entwick-

lungsmilieus entstehen, die zu unterschiedlichen Leistungsentwicklungen von Kindern mit SPF beitragen (Bos et al. 2010) und im Laufe der Schulkarriere eine Akkumulation von Bildungsnachteilen für Kinder aus sozial schwachen Familien nach sich ziehen (Pfahl 2012).

Im Gegensatz dazu wird unter der Annahme einer geschützten Lernumgebung in Förderschulen argumentiert, dass diese den besonderen Bedürfnissen der Schülerinnen und Schüler mit SPF besser gerecht werden können (Markussen 2004; Peetsma et al. 2001; Tent et al. 1991). Demnach wäre es in Förderschulen möglich, die Lehrpläne und Unterrichtsmaterialien so anzupassen, dass sie den Lernvoraussetzungen von Kindern mit SPF entsprechen (Vaughn et al. 1996). Den Lehrkräften könnte durch den geringeren inhaltlichen Umfang der Lehrpläne z. B. mehr Zeit zum Erklären der Lerninhalte zur Verfügung stehen (Schumann 2007). Durch die kleineren Lerngruppen in Förderschulen bestünden zudem bessere Voraussetzungen für eine individualisierte Förderung von Schülerinnen und Schülern mit SPF (Schor 2003; Wocken 2000). Weiterhin würden Kinder mit SPF in der Förderschule einen Schonraum vorfinden (Schumann 2007), der sie vor Leistungsdruck, schulischem Versagen und sozialer Ausgrenzung schütze (Haeberlin 1991; Bear et al. 2002; Preuss-Lausitz 2001; Tent et al. 1991).

4.2.2.2 Empirische Befunde zum Einfluss der Art der Beschulung auf die Schulleistung von Kindern mit SPF

Die Mehrzahl der internationalen Untersuchungen zum Einfluss der Art der Beschulung auf Schulleistungen von Kindern mit SPF haben Vorteile für Regelschulen festgestellt (Baker et al. 1994; Ruijs und Peetsma 2009; Markussen 2004; Myklebust 2006). Ruijs und Peetsma (2009) kommen auf der Grundlage einer Übersicht internationaler Studien zum Schluss, dass der Unterricht von Kindern mit SPF in Regelschulen neutrale oder positive Effekte auf deren schulische Leistungen hat. Peetsma und Kollegen (2001) führten eine Studie durch, in der sie die Leistungsentwicklung von Kindern mit SPF in einer niederländischen Schülerkohorte in Abhängigkeit von der Art ihrer Beschulung im Längsschnitt untersuchten. Dabei wurden an Regel- und Förderschulen Schülerinnen und Schüler mit SPF in der zweiten Jahrgangsstufe mit ähnlichen Ausprägungen in zentralen Hintergrundmerkmalen ausgewählt und anschließend in ihren schulischen Leistungen verglichen. Nach vier Jahren zeigte sich, dass Kinder mit SPF in Regelschulen deutlich höhere schulische Leistungen in mathematischen und unterrichtssprachlichen Kompetenzen aufwiesen als vergleichbare Kinder in Förderschulen.

Auch die Ergebnisse älterer Metaanalysen weisen darauf hin, dass die schulischen Leistungen von Kindern mit SPF in sprachlichen und mathematischen Bereichen in Regelschulen höher ausgeprägt sind als in Förderschulen. Die Größe dieser Effekte variiert dabei von geringen (Baker et al. 1994; Carlberg und Kavale 1980) bis mittelstarken Leistungsvorsprüngen der Schülerinnen und Schüler in Regelschulen (Wang und Baker 1985). In einzelnen Studien zeigen sich allerdings auch neutrale oder sogar negative Effekte integrierter bzw. inklusiver Beschulung für Schülerinnen und Schüler mit SPF (Cole et al. 2004; Harrington 1997; Lindsay 2007; Rogers und Thiery 2003; Zigmond 2003).

Im deutschsprachigen Raum stellte die Forschungsgruppe um Wocken (2000, 2005; Wocken und Gröhlich 2009) fest, dass Schülerinnen und Schüler, die eine Förderschule mit dem Förderschwerpunkt *Lernen* besuchten, selbst noch zu Beginn der 7. Jahrgangsstufe deutlich schwächere Rechtschreibleistungen aufwiesen als Hauptschülerinnen und -schüler der 5. Jahrgangsstufe bzw. Kinder am Ende der Primarstufe mit vergleichbaren kognitiven Grundfähigkeiten. Auch Haeberlin (1991) berichtet für Kinder mit SPF, die in Regelschulen unterrichtet wurden, bessere schulische Leistungen im Vergleich zu Kindern in Förderschulen. Im Gegensatz dazu fanden Lehmann und Hoffmann (2009), dass Jugendliche mit dem Förderschwerpunkt *Lernen* von Unterricht in Förderschulen profitieren können. In der 7. und 8. Jahrgangsstufe waren ihre sprachlichen und mathematischen Leistungen in Regelschulen zwar etwas höher als in Förderschulen, dieses Muster kehrte sich jedoch in der 9. und 10. Klasse zugunsten der Förderschülerinnen und -schüler um.

Der Forschungsstand zum Einfluss der Art der Beschulung auf die schulischen Leistungen von Schülerinnen und Schülern mit SPF ist bislang jedoch nur bedingt belastbar. Wesentliche Limitationen bestehen darin, dass nur selten standardisierte Leistungstests verwendet und adäquate Vergleichsgruppen einbezogen werden. Meistens werden nur wenige Hintergrundmerkmale der untersuchten Schülerinnen und Schüler kontrolliert, so dass die berichteten Effekte z. T. auf bereits vor der Untersuchung bestehende Unterschiede zwischen Kindern mit SPF in Regel- und Förderschulen zurückzuführen sein könnten (Zigmond 2003). Bisherige Studien fokussieren zudem stark auf Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen*, während andere Förderschwerpunkte bislang kaum Beachtung fanden.

4.2.3 Fragestellung und Hypothesen

Ausgehend von den beschriebenen theoretischen Annahmen und empirischen Befunden wird untersucht, ob sich das erreichte Niveau schulischer Leistungen von Kindern mit SPF in Abhängigkeit von der Art ihrer Beschulung unterscheidet. Dabei werden schulische Kompetenzen in sprachlichen und mathematischen Kompetenzbereichen betrachtet. In Übereinstimmung mit vorangegangenen empirischen Untersuchungen wird davon ausgegangen, dass Schülerinnen und Schüler mit SPF an Regelschulen unter Konstanthalten relevanter Hintergrundmerkmale höhere schulische Kompetenzen aufweisen als Kinder in Förderschulen. Da sich die Mehrzahl bisheriger Studien auf den Förderschwerpunkt *Lernen* konzentriert, soll zudem geprüft werden, ob sich die Effekte für verschiedene Förderschwerpunkte nachweisen lassen.

4.2.4 Methoden

4.2.4.1 Stichprobe

Die vorliegende Untersuchung basiert auf den Daten des IQB-Ländervergleichs 2011 für die Primarstufe (Stanat et al. 2012). In dieser querschnittlich angelegten Studie werden Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik mit standardisierten Leistungstests erfasst.

Zur Beantwortung der Fragestellung wurde eine Teilstichprobe von Schülerinnen und Schülern mit einem diagnostizierten SPF und gültigen Leistungswerten in den Fächern Deutsch ($N = 1071$) und Mathematik ($N = 1053$) in die Analysen einbezogen. Diese Kinder konnten anhand von Angaben ihre Schule eindeutig einem der drei Förderschwerpunkte *Lernen*, *Sprache* oder *emotionale und soziale Entwicklung* zugeordnet werden. Zum Zeitpunkt der Erhebung wurden 658 Kinder in 348 Regelschulen und 413 Kinder in 46 Förderschulen unterrichtet (durchschnittliches Alter 10,9 Jahre, $SD = 0,63$; 32,7% Mädchen).⁷ In Tabelle 7 wird die Verteilung der Schülerinnen und Schüler mit SPF an Regel- bzw. Förderschulen nach den betrachteten Fächern und Förderschwerpunkten dargestellt.

⁷ Die berichteten Stichprobengrößen beziehen sich auf das Fach Deutsch. Die nur leicht abweichenden Zahlen für das Fach Mathematik sind in Tabelle 7 dargestellt.

Fehlende Werte in den Schüler- und Elternbefragungen wurden mehrfach imputiert und durch je 15 plausible Werte ersetzt (Schafer 1997; van Buuren und Groothuis-Oudshoorn 2011). Die nachfolgenden Analysen und Auswertungsschritte wurden für die 15 imputierten Datensätze separat durchgeführt und anschließend zusammengefasst (vgl. Rubin 1987).

Tabelle 7. *Stichprobenumfänge im Überblick*

Teilgruppe	Fach	$N_{\text{insgesamt}}$	$N_{\text{Regelschule}}$	$N_{\text{Förderschule}}$
Schülerinnen und Schüler mit SPF ^a (insgesamt)	Deutsch	1071	658	413
	Mathematik	1053	659	394
Schülerinnen und Schüler mit dem Förderschwerpunkt <i>Lernen</i>	Deutsch	536	278	258
	Mathematik	523	282	241
Schülerinnen und Schüler mit dem Förderschwerpunkt <i>Sprache</i>	Deutsch	301	166	135
	Mathematik	301	167	134
Schülerinnen und Schüler mit dem Förderschwerpunkt <i>emotionale und soziale Entwicklung</i>	Deutsch	234	214	20
	Mathematik	229	210	19

^a Sonderpädagogischer Förderbedarf (SPF)

4.2.4.2 Messinstrumente

4.2.4.2.1 Schulische Kompetenzen

Im Ländervergleich 2011 wurden verschiedene Teilkompetenzen in den Fächern Deutsch und Mathematik erhoben. Im Fach Deutsch handelt es sich um Kompetenzen in den Bereichen Lesen und Zuhören (Böhme und Bremerich-Vos 2012), im Fach Mathematik um einen Globalwert, der sich aus fünf inhaltsbezogenen Kompetenzbereichen (Zahlen und Operationen; Raum und Form; Muster und Strukturen; Größen und Messen; sowie Daten, Häufigkeit und Wahrscheinlichkeit) zusammensetzt (Roppelt und Reiss 2012). Die in Regelschulen und Förderschulen eingesetzten Testhefte unterschieden sich in Bezug auf die Anzahl der Aufgaben, den Schwierigkeitsgrad der Aufgaben und die Bearbeitungszeit. In den Testheften für die Förderschulen kamen weniger und leichtere Aufgaben zum Einsatz.⁸ Entsprechend betrug die Bearbeitungszeit an Regelschulen 80 Minuten und an Förderschulen 40 Minuten.

Um die Vergleichbarkeit der Testergebnisse zwischen den betrachteten Schülergruppen zu prüfen, wurden die Testaufgaben für die Teilstichprobe der Kinder in Förderschulen separat frei skaliert und anschließend die Itemfitparameter mit denen für die Gruppe der Regelschüle-

⁸ In Bremen erhielten alle Schülerinnen und Schüler mit SPF unabhängig von der Art der Beschulung die verkürzten Testhefte.

rinnen und -schüler verglichen. Außerdem wurden *Differential Item Functioning*-Analysen zur Beurteilung der Testfairness durchgeführt. Damit sollte sichergestellt werden, dass bei annähernd gleichen Fähigkeiten die Lösungswahrscheinlichkeiten für die einzelnen Testaufgaben für Kinder in Förderschulen und für Kinder in Regelschulen ähnlich ausgeprägt sind. Erste Befunde deuten darauf hin, dass die in Regel- und Förderschulen eingesetzten Testaufgaben ähnlich gute psychometrische Eigenschaften aufweisen (Kuhl et al. 2013).

4.2.4.2.2 Kognitive Grundfähigkeiten

Die Erfassung der kognitiven Grundfähigkeiten der Schülerinnen und Schüler basiert auf zwei Untertests des Kognitiven Fähigkeitstests (KFT 4-12+R, Heller und Perleth, 2000): dem Untertest *Wortschatz* als Indikator für die verbalen kognitiven Grundfähigkeiten und dem Untertest *Figurenanalogien* als Indikator für die räumlichen kognitiven Grundfähigkeiten. Beide Skalen können zur Kontrolle des Vorwissens und der damit verbundenen differenziellen Eingangsselektivität von Regel- und Förderschulen eingesetzt werden (Baumert et al. 2006).

4.2.4.2.3 Sozio-kultureller Hintergrund

Als Kontrollvariablen für die Schätzung des Einflusses der Art der Beschulung auf die Schulleistung dienen die sozialen, kulturellen und zuwanderungsbezogenen Hintergrundmerkmale der Schülerinnen und Schüler. Die Beschreibung des sozialen Status der Familien erfolgte über den jeweils höchsten *International Socio-Economic Index of Occupational Status* (ISEI) des Vaters bzw. der Mutter (Ganzeboom et al. 1992). Das Bildungsniveau der Familien wurde über den höchsten Bildungsabschluss der Eltern erfasst (*International Standard Classification of Education*, ISCED; OECD, 1999). Weiterhin gingen Schülerangaben zum häuslichen Buchbestand und die von Eltern geäußerten Bildungsaspirationen (für das Kind gewünschter bzw. erwarteter Schulabschluss) in die Analysen ein. Zudem wurden Elterneinschätzungen zur Bedeutsamkeit der Vermittlung bestimmter Erziehungsziele in der Familie erfasst. Dabei sollten die Eltern einschätzen, für wie wichtig sie die Vermittlung solcher Merkmale wie persönliche Selbstständigkeit, Ordnung und Disziplin, Bereitschaft zum Lernen und solide Schulkenntnisse in der Familie erachten (Baumert et al. 2008). Hohe Werte auf der durch Aggregation der 18 Items gebildeten Skala (Cronbachs $\alpha = .90$) weisen darauf hin, dass Eltern diesen Erziehungszielen eine hohe Bedeutung beimessen. Der Zuwanderungshintergrund der Schülerinnen und Schüler wurde anhand des Geburtslands der Eltern erfasst (mindestens ein Elternteil nicht in Deutschland geboren). Zusätzlich wurde die

in der Familie gesprochene Sprache berücksichtigt und unterschieden, ob in der Familie überwiegend Deutsch oder überwiegend eine andere Sprache gesprochen wird.

4.2.4.2.4 Lerngelegenheiten und Bundesland

Zu den betrachteten Indikatoren für institutionelle Lerngelegenheiten, die Kinder mit SPF vor Eintritt in eine Regel- oder Förderschule hatten, gehören das Schuleintrittsalter und die Dauer des Besuchs einer Kindertagesstätte bzw. Vorschule. Überdies wurde aufgrund der erheblichen Unterschiede zwischen den Ländern in Bezug auf den Anteil von Kindern mit einem diagnostizierten SPF und deren Verteilung auf Regel- und Förderschulen (Dietze 2012; Stanat et al. 2012) das Bundesland als Kontrollvariable aufgenommen.

4.2.4.2.5 Diagnostizierte Teilleistungsstörung

Neben dem Förderschwerpunkt wurden auch diagnostizierte Teilleistungsstörungen der Kinder erfasst und in den Auswertungen als Kontrollvariable berücksichtigt. Diese umfassten die Kategorien (1) Aufmerksamkeitsdefizit-Hyperaktivitätsstörung (ADHS), (2) Lernstörungen (isolierte Rechtschreibstörung, Lese-Rechtschreibstörung, Rechenstörung), (3) Mehrfachdiagnose (Kombination aus ADHS und einer Lernstörung) sowie (4) nicht definierte Teilleistungsstörungen.

4.2.4.3 Datenanalyse

Ziel der vorliegenden Untersuchung ist es, den Effekt der Art der Beschulung (Regelschule vs. Förderschule) auf die Schulleistung von Grundschulkindern mit SPF zu bestimmen. Eine optimale Bedingung für einen validen Vergleich zwischen Regel- und Förderschulen bestünde in einer Zuweisung der Kinder zu einer der beiden Schularten, die unabhängig von den schulbezogenen Kompetenzen der Kinder erfolgt. In experimentellen Versuchsplänen wird dies durch den Prozess der Randomisierung erreicht. Im vorliegenden Fall ist es jedoch nicht möglich, Schülerinnen und Schüler zufällig einer Regel- bzw. Förderschule zuzuweisen. Unterschiede in der Schulleistung können daher nicht eindeutig auf die Art der Beschulung attribuiert werden, da diese auch durch Unterschiede zwischen den betrachteten Gruppen verursacht worden sein könnten, die bereits vor der Zuweisung zu Regel- und Förderschulen bestanden (Morgan und Winship 2007).

Propensity Score Matching-Verfahren können zur statistischen Kontrolle solcher Selektionseffekte eingesetzt werden, um ggf. identifizierte Leistungsunterschiede eindeutiger auf den

Gruppenfaktor, also hier die Art der Beschulung, zurückführen zu können (Rosenbaum und Rubin 1983). Das Grundprinzip beim Propensity Score Matching besteht darin, für jede Schülerin und jeden Schüler der Stichprobe anhand der für die Zuweisung relevanten Merkmale den Wahrscheinlichkeitswert – den so genannten Propensity Score – zu berechnen, in eine der beiden Schularten überzugehen. Anschließend werden nur diejenigen Schülerinnen und Schüler beider Gruppen in ihrer Leistung verglichen, die ähnliche Propensity Scores aufweisen. Die Anwendung des Propensity Score Matchings erfolgte in der vorliegenden Untersuchung in fünf Schritten (vgl. Becker 2011):

Im ersten Schritt wurde anhand einer multivariaten Varianzanalyse und *Chi*²-Tests geprüft, ob sich Schülerinnen und Schüler mit SPF, die eine Regelschule besuchen, systematisch von Schülerinnen und Schülern an Förderschulen unterscheiden. Ziel dieser Analysen war es, Merkmale zu identifizieren, die sowohl mit der Zuweisung zur Regel- bzw. Förderschule als auch mit den erreichten Kompetenzständen in Deutsch und Mathematik zusammenhängen. Im zweiten Schritt wurde auf Basis der so definierten Kontrollvariablen die bedingte Wahrscheinlichkeit (Propensity Score) des Besuchs einer Regel- oder Förderschule mit Hilfe von binären logistischen Regressionen ermittelt.⁹ Im dritten Schritt, dem eigentlichen Matching, wurden Schülerinnen und Schüler mit SPF in Regel- und Förderschulen auf Basis eines *Full Matching*-Verfahrens ausgewählt, die ähnliche Propensity Scores aufweisen. Das Propensity Score Matching wurde mit dem Paket MatchIt (Ho et al. 2011) im Statistikprogramm R (R Development Core Team 2010) durchgeführt. In einem vierten Schritt wurde evaluiert, inwieweit das Matching zu homogenen Untersuchungsgruppen geführt hat. Ein wichtiges Mittel zur Bestimmung der durch das Matching erreichten Balance ist die Betrachtung der standardisierten Differenzen (standardisierter Bias; Rosenbaum & Rubin 1985) zwischen den untersuchten Gruppen (Ho et al. 2007). Je geringer diese ausfallen, desto stärker ähneln sich beide Untersuchungsgruppen nach dem Matching. Weiterhin wurde auch die Verteilung der Propensity Scores und der Kontrollvariablen zwischen den Untersuchungsgruppen berücksichtigt.

Wenn durch den vorherigen Analyseschritt sichergestellt werden konnte, dass sich die Untersuchungsgruppen nach dem Matching hinreichend ähneln, können anschließend in einem fünften Schritt inferenzstatistische Auswertungsverfahren mit der gematchten Stich-

⁹ Aufgrund der günstigeren Verteilungseigenschaften wurde der vorhergesagte Propensity Score Logit-transformiert (Rosenbaum und Rubin 1985).

probe durchgeführt werden. Da sich beide Untersuchungsgruppen nach dem Matching im Idealfall nur noch durch die Art ihrer Beschulung und nicht mehr in Bezug auf andere leistungsdeterminierende Merkmale unterscheiden, können Schulleistungsdifferenzen mit höherer Plausibilität auf die Art der Beschulung zurückgeführt werden. Allerdings sind Aussagen zur Wirkung der Beschulungsart auf die Schulleistungen nur hinsichtlich der einbezogenen Kontrollvariablen und dem Überlappungsbereich in der Ausprägung der Kontrollvariablen zwischen beiden Gruppen gültig. Deshalb ist es wichtig, dass sich die Propensity Score-Verteilungen der zu vergleichenden Gruppen ausreichend überlappen (*Area of Common Support*).

Im Anschluss an das Propensity Score Matching wurde der *Average Treatment Effect on the Treated (ATT)* berechnet. Die Bestimmung des ATT basiert auf dem Kausalmodell von Rubin (1974) und indiziert in der vorliegenden Untersuchung die durchschnittliche Wirkung des Besuchs einer Regelschule relativ zum Besuch einer Förderschule auf die Leistungen der Schülerinnen und Schüler mit SPF in Regelschulen. Die Schätzung des ATTs erfolgte unter Verwendung eines linearen Regressionsmodells, bei dem für Unterschiede in kognitiven Grundfähigkeiten und im sozialen Status kontrolliert wurde. Dadurch sollte so weit wie möglich gewährleistet werden, dass nach dem Matching ggf. weiterhin bestehende Unterschiede zwischen beiden Untersuchungsgruppen in den ausgewählten Kovariaten nicht mit der Effektschätzung konfundiert sind (Stuart 2010). Die Schätzung des ATTs wurde mit dem Programmpaket Zelig in R (Imai et al. 2008) durchgeführt.

Eine Limitation bei Anwendung des Propensity Score Matchings in der vorliegenden Studie stellt das querschnittliche Untersuchungsdesign dar. So kann nicht ausgeschlossen werden, dass bestimmte Merkmale (z.B. kognitive Grundfähigkeiten), die im Rahmen des Matchings berücksichtigt wurden, ebenfalls durch die Beschulungsart beeinflusst wurden (*Posttreatment Selection Bias*, Schafer und Kang 2008). Dies ist insofern problematisch, als dass die Zuweisung zu einer Regel- oder Förderschule selbst nach dem Matching noch mit den erzielten Schulleistungen konfundiert sein könnte. Folglich kann der Einfluss der Beschulungsart auf die Schulleistungen von Kindern mit SPF nicht vollkommen verzerrungsfrei geschätzt werden. Die resultierenden Ergebnisse sind daher mit entsprechender Unsicherheit behaftet.

4.2.5 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Datenanalyse anhand der oben beschriebenen fünf Schritte des Vorgehens beim Propensity Score Matching dargestellt.

4.2.5.1 Unterschiede zwischen Kindern mit SPF in Abhängigkeit von der Beschulungsart

Im ersten Schritt wurde anhand einer multivariaten Varianzanalyse mit dem ersten imputierten Datensatz geprüft, in welchen Merkmalen sich Kinder mit SPF in Abhängigkeit von ihrer Beschulungsart systematisch voneinander unterscheiden (Tab. 8). Nahezu alle Gruppenunterschiede waren unter Kontrolle für Mehrfachtestung (Bonferroni-Korrektur) signifikant. Die deutlichsten Differenzen ergaben sich für die kognitiven Grundfähigkeiten, die Bildungsaspirationen der Eltern und den sozialen Status der Familien. Kinder mit SPF, die eine Regelschule besuchten, zeigten deutlich höhere kognitive Grundfähigkeiten, sowohl im verbalen ($d = 0,80$) als auch im räumlichen Bereich ($d = 0,49$), als Kinder in Förderschulen. Weiterhin gaben Eltern von Kindern mit SPF in Regelschulen höhere Bildungsaspirationen an als Eltern, deren Kinder in einer Förderschule unterrichtet wurden. Auch der soziale Status von Kindern mit SPF in Regelschulen war verglichen mit Kindern in Förderschulen im Durchschnitt höher ($d = 0,36$). Keine statistisch bedeutsamen Unterschiede zwischen den Untersuchungsgruppen fanden sich lediglich für das Alter und die Vermittlung von Erziehungszielen in der Familie.

Tabelle 8. *Unterschiede zwischen Kindern mit SPF nach Art der Beschulung*

Kovariaten	Förderschule		Regelschule		<i>F</i>	<i>d</i> ^a	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Kognitive Grundfähigkeiten: verbal	9.04	4.99	12.72	4.33	162.83	0.80	<.001
Kognitive Grundfähigkeiten: räumlich	10.22	6.18	13.45	6.91	60.17	0.49	<.001
Alter (in Jahren)	10.97	0.61	10.86	0.64	8.73	−0.19	<.01
Sozialer Status (HISEI) ^b	36.79	13.27	41.88	14.57	33.21	0.36	<.001
Höchster Bildungsabschluss in der Familie (HISCED) ^c	3.10	1.13	3.57	1.20	40.77	0.40	<.001
Gewünschter Schulabschluss: Hauptschule	4.22	1.90	3.26	2.01	60.59	−0.49	<.001
Voraussichtlicher Schulabschluss: mindestens Realschulabschluss	2.75	0.90	3.20	0.78	73.38	0.54	<.001
Bücher zu Hause	2.51	1.24	2.82	1.23	15.96	0.25	<.001
Vermittlung von Erziehungszielen	3.20	0.52	3.20	0.48	0.00	0.00	.96
Schuleintrittsalter (gruppiert)	2.48	0.73	2.20	0.65	42.67	−0.41	<.001

N = 1071. ^aCohens *d*, ^bHighest international socio-economic index of occupational status,

^cHighest international standard classification of education

4.2.5.2 Vorhersage des Besuchs einer Regelschule

Im nächsten Schritt wurde mit logistischen Regressionsanalysen geprüft, wie gut die ausgewählten Kovariaten geeignet sind, den Besuch einer Regelschule vorherzusagen. Dazu wurden für die Teilstichproben Deutsch und Mathematik separate logistische Regressionsanalysen durchgeführt, da sich die Stichprobenumfänge zwischen beiden Bereichen leicht unterscheiden. Die Ergebnisse der logistischen Regressionsanalyse im Fach Deutsch (siehe Tab. 9) korrespondiert zum Großteil mit den deskriptiven Befunden und Ergebnissen der multivariaten Varianzanalyse. Bei der Schülerschaft mit SPF, die an einer Regelschule unterrichtet wurde, handelte es sich demnach um eine hinsichtlich kognitiver und sozio-kultureller Merkmale selektive Gruppe. Sowohl Kinder mit höheren verbalen kognitiven Grundfähigkeiten als auch Kinder aus bildungsnahen Familien (erfasst anhand des sozialen Status der Familie, den höchsten familiären Bildungsabschluss und die Bildungsaspirationen der Eltern) wiesen eine höhere Wahrscheinlichkeit des Besuchs einer Regelschule auf. Diese Befunde weisen auf eine differenzielle Eingangsselektivität der betrachteten Schularten hin. Weiterhin wurden Kinder mit einer Lernstörung eher an Regelschulen unterrichtet als Kinder mit einer anderen oder keiner berichteten Teilleistungsstörung. Dies könnte darauf zurückzu-

führen sein, dass die Diagnose eines Förderbedarfs im Schwerpunkt *Lernen* in den letzten 10 Jahren zurückgegangen ist und eine Tendenz besteht, diese vermehrt in Regelschulen zu unterrichten (Autorengruppe Bildungsberichterstattung 2010; Dietze 2011). Außerdem zeigte sich, dass Kinder mit SPF an Regelschulen beim Schuleintritt jünger waren und kürzer einen Kindergarten besuchten als Kinder an Förderschulen. Eine mögliche Erklärung für den negativen Zusammenhang zwischen Dauer des Kindergartenbesuchs und dem Besuch einer Regelschule könnte darin liegen, dass v.a. Kinder mit einem stärkeren Beeinträchtigungsgrad länger im Kindergarten verweilen, um mit zusätzlichen frühkindlichen Fördermaßnahmen eine Einschulung in eine Förderschule zu verhindern (Morgan et al. 2010; Sullivan & Field 2013). Die Vorhersage des Besuchs einer Regelschule anstelle einer Förderschule durch die logistische Regression gelang sowohl für die Deutsch- als auch die Mathematikstichprobe zuverlässig. In beiden Teilstichproben lag die Varianzaufklärung bei $R^2 = 0,45$ (Nagelkerke 1991).

Tabelle 9. *Ergebnisse der logistischen Regression zur Vorhersage des Besuchs einer Regelschule (Basis: Stichprobe mit Leistungswerten im Fach Deutsch)*

	<i>B (SE)</i>	<i>p</i> -Wert	<i>adj. OR^a</i>	2.5%	97.5%
Kognitive Grundfähigkeiten: verbal	0.84 (0.10)	<.001	2.32	1.90	2.84
Kognitive Grundfähigkeiten: räumlich	0.17 (0.09)	.06	1.18	0.99	1.41
Teilleistungsstörung ^b : ADHS ^c	0.16 (0.29)	.58	1.17	0.67	2.06
Teilleistungsstörung: Lernstörung	1.34 (0.24)	<.001	3.83	2.41	6.10
Teilleistungsstörung: Mischform	0.35 (1.27)	.79	1.41	0.12	17.21
Teilleistungsstörung: undefiniert	0.16 (0.41)	.71	1.17	0.52	2.63
Geschlecht: männlich	-0.46 (0.17)	<.01	0.63	0.45	0.89
Alter	0.06 (0.14)	.68	1.06	0.80	1.40
Sozialer Status (HISEI) ^d	0.20 (0.09)	<.05	1.22	1.01	1.46
Höchster Bildungsabschluss in der Familie (HISCED) ^e	0.18 (0.09)	<.05	1.20	1.01	1.42
Gewünschter Schulabschluss: Hauptschule	-0.21 (0.09)	<.05	0.81	0.68	0.96
Voraussichtlicher Schulabschluss: mindestens Realschulabschluss	0.23 (0.09)	<.01	1.26	1.06	1.49
Bücher zu Hause	-0.01 (0.09)	.91	0.99	0.83	1.17
Vermittlung von Erziehungszielen	-0.08 (0.08)	.31	0.92	0.79	1.08
Zuwanderungshintergrund (0/1) ^f	-0.15 (0.22)	.49	0.86	0.56	1.32
Deutsch zu Hause (0/1) ^g	0.01 (0.33)	.97	1.01	0.53	1.93
Schuleintrittsalter	-0.25 (0.09)	<.01	0.78	0.66	0.93
Besuch Kindergarten ^h : <1 Jahr	-1.75 (0.70)	<.05	0.17	0.04	0.69
Besuch Kindergarten: 1-2 Jahre	-1.10 (0.56)	<.05	0.33	0.11	0.98
Besuch Kindergarten: 2-3 Jahre	-1.13 (0.49)	<.05	0.32	0.12	0.84
Besuch Kindergarten: > 3 Jahre	-0.97 (0.47)	<.05	0.38	0.15	0.95
<i>R</i> ² (Nagelkerke)	.45				

N = 1071. ^a Adjustierte odds ratios, ^b Referenzkategorie: keine Teilleistungsstörung bzw. fehlende Angabe, ^c Aufmerksamkeitsdefizit-Hyperaktivitätsstörung, ^d Highest international socio-economic index of occupational status, ^e Highest international standard classification of education, ^f 1 = beide Eltern in Deutschland geboren, ^g 1 = immer oder fast immer Deutsch, ^h Referenzkategorie: kein Besuch eines Kindergartens

4.2.5.3 Verwendetes Matchingverfahren

Die Auswahl vergleichbarer Kinder mit SPF erfolgte über ein *Full Matching*-Verfahren (Hansen 2004; Rosenbaum 1991). Dabei wurde die Stichprobe auf Basis des Propensity Scores in verschiedene Untergruppen aufgeteilt. Jede Untergruppe bestand aus mindestens einem Kind aus einer Regel- und einer Förderschule, wobei das Verhältnis zwischen den Schülerinnen und Schülern beider Schularten über die verschiedenen Untergruppen variierte. Die Anzahl der Untergruppen wurde iterativ so bestimmt, dass sich die Unterschiede auf den Propensity Scores in jeder Untergruppe minimierten. Die anschließende Schätzung des ATTs erfolgte unter Verwendung von Gewichtungsfaktoren, die die Anzahl der Schülerinnen und Schüler pro Untergruppe widerspiegeln. Schülerinnen und Schüler mit extremen Ausprägungen auf dem Propensity Score wurden von den anschließenden Analysen ausgeschlossen.

4.2.5.4 Vergleich der Untersuchungsgruppen nach dem Matching

In Tabelle 10 werden die Verteilungen der Propensity Scores vor und nach dem Matching sowie der standardisierte Bias abgebildet. Die Darstellung erfolgt dabei sowohl für die gesamte Stichprobe als auch separat für die Förderschwerpunkte *Lernen* und *Sprache*. Aufgrund der geringen Anzahl von Schülerinnen und Schülern in Förderschulen mit Schwerpunkt *emotionale und soziale Entwicklung* konnte für diese Teilstichprobe kein Matching durchgeführt werden.

Tabelle 10. *Verteilung der Propensity Scores vor und nach dem Matching und standardisierter Bias (SB)*

Teilgruppe	Schulart	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	Schiefe	Exzess	<i>SB</i> ^a
Alle Kinder mit SPF ^b									
Vor dem Matching	FS ^c	413	-4.92	4.38	-0.65	1.53	0.12	-0.10	1.47
	RS ^d	658	-2.71	6.64	1.57	1.48	0.13	0.10	
Full Matching	FS	389	-3.09	4.16	1.39	1.30	-0.23	0.01	-0.01
	RS	622	-3.10	4.14	1.38	1.29	-0.28	-0.02	
Kinder mit dem Förder- schwerpunkt <i>Lernen</i>									
Vor dem Matching	FS	258	-6.61	4.02	-1.82	1.95	0.13	-0.05	1.20
	RS	278	-2.91	24.72	3.40	5.81	2.61	5.88	
Full Matching	FS	148	-2.51	4.99	1.81	1.75	-0.14	-0.54	-0.02
	RS	234	-2.58	4.95	1.78	1.70	-0.24	-0.58	
Kinder mit dem Förder- schwerpunkt <i>Sprache</i>									
Vor dem Matching	FS	135	-18.16	3.48	-1.37	2.17	-3.66	26.03	1.10
	RS	166	-2.35	36.04	7.35	11.00	1.29	0.21	
Full Matching	FS	110	-2.57	2.08	0.27	1.09	-0.52	-0.03	-0.01
	RS	85	-2.81	2.08	0.26	1.09	-0.64	0.26	

^a Standardisierter Bias, ^b Sonderpädagogischer Förderbedarf, ^c Förderschule, ^d Regelschule

Es zeigte sich, dass das Full Matching zu einer starken Reduktion des standardisierten Bias auf dem Propensity Score führte. Auch die Verteilung der Propensity Scores und die Unterschiede auf den einzelnen Kovariaten weisen darauf hin, dass eine Parallelisierung der Untersuchungsgruppen erreicht werden konnte (Abb.7). Somit kann angenommen werden, dass sich Kinder in Regel- und Förderschulen nach dem Matching hinsichtlich der betrachteten Kovariaten nicht mehr bedeutsam unterscheiden. Zudem liegt in der gematchten Stichprobe eine hinreichend große Überlappung der Propensity Scores zwischen den Untersuchungsgruppen vor.

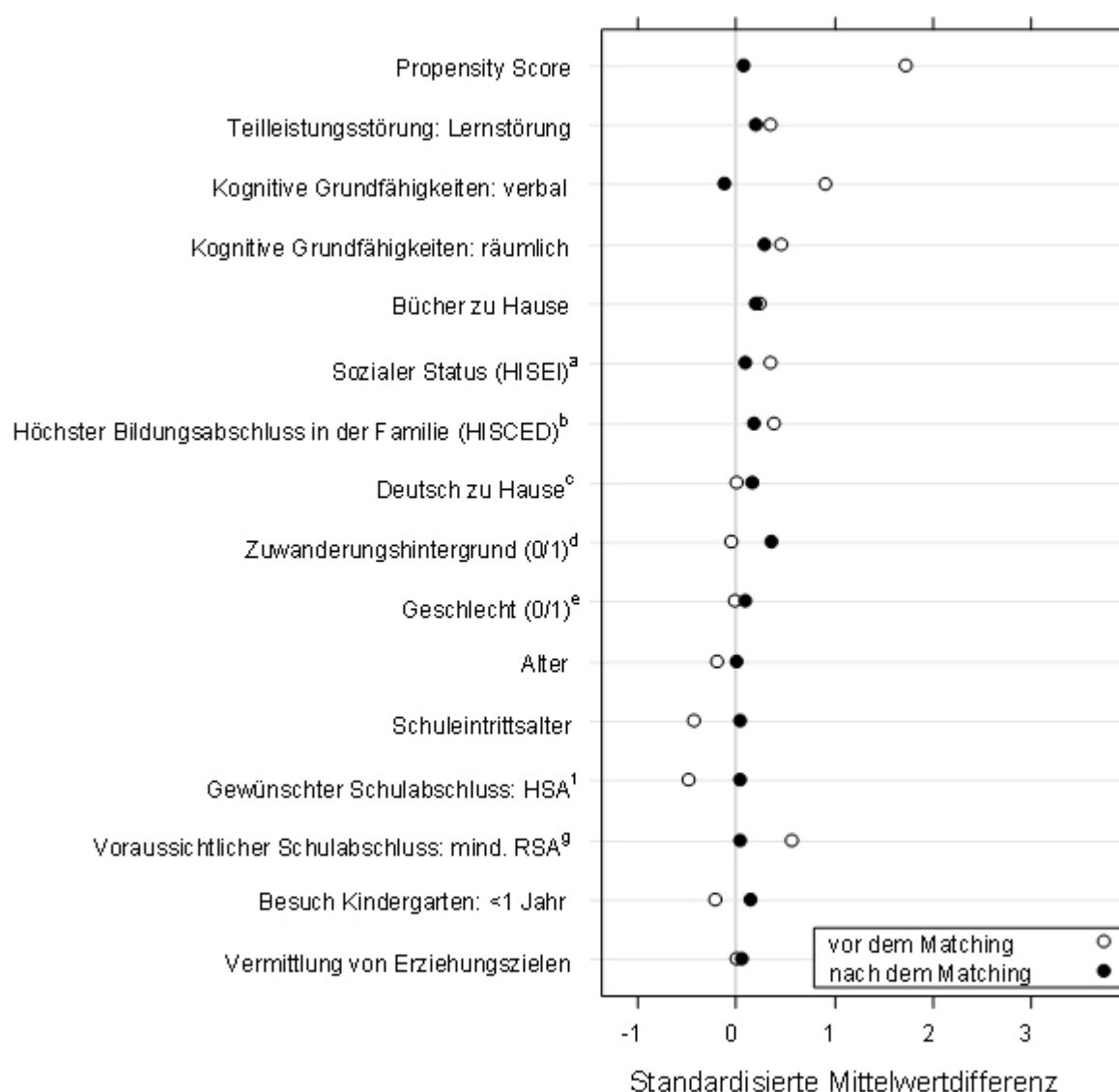


Abbildung 7. Standardisierte Mittelwertunterschiede zwischen Kindern mit SPF in Regel- und Förderschulen vor und nach dem Full Matching.

Anmerkungen: Die Darstellung bezieht sich auf die gesamte Stichprobe der Kinder mit sonderpädagogischem Förderbedarf, für die Leistungswerte im Fach Deutsch vorliegen ($N = 1071$).

^a Highest international socio-economic index of occupational status, ^b Highest international standard classification of education, ^c 1 = immer oder fast immer Deutsch, ^d 1 = beide Eltern in Deutschland geboren, ^e 1 = männlich., ^f Hauptschulabschluss, ^g Realschulabschluss

4.2.5.5 Schätzung differenzieller Effekte der Beschulungsart auf die Schulleistung

Im abschließenden Auswertungsschritt wurden die Effekte der Art der Beschulung auf die Schulleistungen in Deutsch (Lesen und Zuhören) und Mathematik innerhalb der gematchten Stichproben geschätzt. In Tabelle 11 werden die ATTs (s. Abschnitt 4.3) und Standardfehler für die aggregierten Kompetenzwerte berichtet. Die Mittelwertdifferenzen wurden an der Standardabweichung der gesamten Stichprobe des IQB-Ländervergleichs Primarstufe 2011 normiert ($M = 500$, $SD = 100$) und lassen sich im Rahmen der Konventionen von Cohen (1992) als Effektstärken interpretieren (Cohens d).¹⁰

Tabelle 11. *Effekte des Besuchs einer Regelschule auf die Kompetenzen in Deutsch und Mathematik nach dem Full Matching*

Teilgruppe	Deutsch Lesen			Deutsch Zuhören			Mathematik		
	ATT ^a	SE ^b	<i>p</i>	ATT	SE	<i>p</i>	ATT	SE	<i>p</i>
Alle Kinder mit SPF ^c	37.57	5.95	<.001	50.00	6.78	<.001	43.03	6.34	<.001
Kinder mit dem Förderschwerpunkt <i>Lernen</i>	48.62	9.72	<.001	75.19	9.74	<.001	79.53	8.37	<.001
Kinder mit dem Förderschwerpunkt <i>Sprache</i>	26.90	8.77	<.01	48.88	10.56	<.001	6.83	8.40	.42

^a Average treatment effect on the treated (gemittelt); Skala mit Mittelwert 500 und Standardabweichung 100; positive Werte zugunsten von Kindern mit SPF in Regelschulen

^b Gepoolter Standardfehler (Rubin 1987), ^c Sonderpädagogischer Förderbedarf mit Schwerpunkt *Lernen*, *Sprache* oder *emotionale und soziale Entwicklung*

Im Hinblick auf die Lesekompetenzen zeigte sich, dass Schülerinnen und Schüler mit SPF in Regelschulen insgesamt signifikante Leistungsvorsprünge im Vergleich zur gematchten Schülergruppe in Förderschulen erzielten. Der korrespondierende ATT betrug 38 Punkte ($p < 0,001$), was nach Cohen (1992) einem geringen Effekt zugunsten der Kinder mit SPF in Regelschulen entspricht. Eine Differenzierung nach Förderschwerpunkten ergab, dass insbesondere Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* vom gemeinsa-

¹⁰ Die Effektstärke Cohens d quantifiziert den standardisierten Mittelwertunterschied zwischen Kindern mit SPF in Regel- und Förderschulen relativiert an ihrer gemeinsamen Standardabweichung. Zur Beurteilung der praktischen Relevanz der Effekte kann die folgende Klassifizierung herangezogen werden (vgl. Cohen 1992): Ein kleiner Effekt entspricht $d = 0,20$ Standardabweichungen (übertragen auf die vorliegende Untersuchung 20 Punkte), ein mittlerer Effekt $d = 0,50$ Standardabweichungen (50 Punkte auf der Metrik der eingesetzten Kompetenztests) und ein großer Effekt $d = 0,80$ (80 Punkte).

men Unterricht in Regelschulen profitierten. Hier verzeichneten Kinder mit SPF, die in einer Regelschule unterrichtet wurden, signifikant bessere Lesekompetenzen als vergleichbare Kinder in Förderschulen ($ATT = 49$ Punkte, $p < 0,001$). Deutlich geringer fiel der Unterschied für den Förderschwerpunkt *Sprache* aus. Hier zeigte sich ein durchschnittlicher Leistungsvorsprung der Kinder in Regelschulen von $ATT = 27$ Punkten ($p < 0,01$) gegenüber vergleichbaren Kindern in Förderschulen.

Der differenzielle Befund im Hinblick auf den Förderschwerpunkt zeichnete sich auch für den Kompetenzbereich Zuhören ab. Der globale ATT über alle betrachteten Förderschwerpunkte lag bei durchschnittlich 50 Punkten ($p < 0,001$) zugunsten des Besuchs einer Regelschule. Wiederum erwiesen sich die Leistungsunterschiede für Kinder mit dem Förderschwerpunkt *Lernen* als besonders ausgeprägt. Hier betrug der Leistungsvorsprung von Schülerinnen und Schülern in Regelschulen 75 Punkte ($p < 0,001$) gegenüber vergleichbaren Kindern in Förderschulen. Im Förderschwerpunkt *Sprache* waren die Leistungsunterschiede geringer und lagen im Mittel bei $ATT = 49$ Punkten ($p < 0,001$) zugunsten der Kinder mit SPF in Regelschulen.

Für das Fach Mathematik zeigte sich über alle Förderschwerpunkte betrachtet ebenfalls eine höhere durchschnittliche Testleistung der Schülerinnen und Schüler mit SPF, die in einer Regelschule unterrichtet wurden ($ATT = 43$ Punkte, $p < 0,001$). Dieser globale Effekt ließ sich vorrangig auf ausgeprägte Leistungsunterschiede der Schülerschaft mit dem Förderschwerpunkt *Lernen* zurückführen. Innerhalb dieser Gruppe wiesen Schülerinnen und Schüler in Regelschulen deutlich höhere mathematische Kompetenzen auf als vergleichbare Schülerinnen und Schüler in Förderschulen ($ATT = 80$ Punkte, $p < 0,001$). Innerhalb der Teilstichprobe des Förderschwerpunktes *Sprache* zeigten sich hingegen keine signifikanten Leistungsunterschiede zwischen Kindern in Regel- und Förderschulen ($ATT = 7$ Punkte, $p = 0,42$).

4.2.6 Zusammenfassung und Diskussion

4.2.6.1 Zusammenfassung der vorliegenden Studie

Ziel der Studie war es, den Einfluss der Art der Beschulung auf die schulischen Kompetenzen von Kindern mit SPF im Bereich Lesen, Zuhören und Mathematik am Ende der vierten Jahrgangsstufe zu untersuchen. Hierzu wurden Daten der IQB-Ländervergleichsstudie 2011 für die Primarstufe herangezogen, in der auch Schülerinnen und Schüler der drei Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* getestet wurden. Um Gruppenunterschiede auf Merkmalen zu kontrollieren, die sowohl für die Zuweisung zu einer Regel- bzw. Förderschule als auch für die Schülerkompetenzen relevant sein könnten, wurden Propensity Score Matching-Verfahren eingesetzt (Rosenbaum und Rubin 1983).

Als Ergebnis kann festgehalten werden, dass Schülerinnen und Schüler mit SPF, die in einer Regelschule unterrichtet wurden, in allen untersuchten Bereichen höhere Leistungen aufwiesen als vergleichbare Schülerinnen und Schüler mit SPF in Förderschulen. Nach Cohen (1992) handelt es sich bei den ermittelten Leistungsunterschieden je nach betrachtetem Kompetenzbereich um geringe bis mittlere Effekte zugunsten des Besuchs einer Regelschule.

Eine nach Förderschwerpunkten differenzierte Betrachtung ergab, dass insbesondere Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* vom Unterricht in Regelschulen zu profitieren scheinen. Für Kinder mit dem Förderschwerpunkt *Sprache* waren die Leistungsunterschiede in den sprachlichen Kompetenzen in Abhängigkeit von der Art der Beschulung hingegen geringer ausgeprägt. Für das Fach Mathematik konnte kein bedeutsamer Leistungsunterschied zwischen Regel- und Förderschülerinnen und -schülern mit dem Förderschwerpunkt *Sprache* nachgewiesen werden.

Um eine inhaltliche Interpretation der Effekte zu ermöglichen, können diese mit der zu erwartenden durchschnittlichen Leistungszunahme zwischen der dritten und vierten Jahrgangsstufe verglichen werden (vgl. Granzer et al. 2009; Stanat et al. 2012). Vorangegangene Large-Scale-Assessments im Primarbereich fanden bezogen auf die Metrik der Bildungsstandards einen mittleren Kompetenzzuwachs von der dritten zur vierten Jahrgangsstufe von circa 60 Punkten in den Kompetenzbereichen Lesen und Zuhören (Behrens et al. 2009; Bremerich-Vos und Böhme 2009) und etwa 80 Punkten in Mathematik für die gesamte Schülerpopulation (Reiss und Winkelmann 2009).

Damit entsprechen die Leistungsvorsprünge der Kinder mit SPF in Regelschulen über alle Förderschwerpunkte hinweg dem Lernzuwachs von etwa einem halben Schuljahr im Lesen und in Mathematik sowie annähernd einem Schuljahr im Zuhören. Für Kinder mit dem Förderschwerpunkt *Lernen* war der Vorsprung in Regelschulen im Vergleich zu Förderschulen noch deutlich größer: Im Kompetenzbereich Lesen und in Mathematik nähert sich der in dieser Studie identifizierte Unterschied einem Leistungsvorsprung von einem Schuljahr; im Kompetenzbereich Zuhören lag er bei mehr als einem Schuljahr zugunsten der Kinder in Regelschulen. Für Schülerinnen und Schüler mit dem Förderschwerpunkt *Sprache* waren dagegen deutlich geringere Leistungsunterschiede zu verzeichnen; sie entsprechen im Kompetenzbereich Lesen aber dennoch knapp einem halben Schuljahr und im Zuhören unter einem Schuljahr zugunsten der Kinder in Regelschulen. In Mathematik zeigten sich für diese Schülergruppe keine bedeutsamen Leistungsunterschiede.

Bei dieser Interpretation der Leistungsunterschiede muss allerdings einschränkend beachtet werden, dass die Schätzungen der Lernzuwächse nicht anhand von Leistungsentwicklungen einzelner Schülerinnen und Schüler sondern durch querschnittliche Vergleiche von Alterskohorten ermittelt wurden. Zudem basieren die Schätzungen der Lernzuwächse auf Daten der Gesamtpopulation von Schülerinnen und Schülern und es ist unklar, inwieweit sie auch für Kinder mit SPF gültig sind. Untersuchungen zum Matthäuseffekt (Walberg und Tsai 1983) legen die Vermutung nahe, dass Kinder mit SPF aufgrund niedrigerer Ausgangsleistungen geringere Lernzuwächse erzielen als Kinder ohne SPF, so dass der Unterschied tendenziell noch größer ausfiele. Allerdings ist auch diese Vermutung bislang ungeprüft (s. aber Peetsma et al. 2001).

Die Befunde der vorliegenden Studie stehen in Einklang mit nationalen und internationalen Untersuchungen zu den Auswirkungen der Beschulungsart auf die schulische Kompetenzentwicklung von Schülerinnen und Schülern mit SPF (Haeberlin 1991; Ruijs und Peetsma 2009; Wocken 2000, 2005; Wocken und Gröhlich 2009). Sie erweitern den Forschungsstand dahingehend, dass eine bundesweite Stichprobe berücksichtigt und eine differenzierte Betrachtung nach Förderschwerpunkten vorgenommen wurde. Eine Stärke der vorliegenden Analysen liegt in der Verwendung von Propensity Score Matching-Verfahren. Dies ermöglichte eine Parallelisierung der Vergleichsgruppen hinsichtlich einer Vielzahl von Schülermerkmalen und somit eine für Analysen mit Querschnittsdaten vergleichsweise genaue Schätzung des Einflusses der Beschulungsart auf die schulischen Kompetenzen. Außerdem können durch das mehrstufige Auswertungsverfahren Aussagen über die differenzielle

Eingangsselektivität an Regel- und Förderschulen getroffen werden. Anhand der logistischen Regression (s. Tab. 9) zeigte sich, dass Kinder mit SPF an Regelschulen einen günstigeren sozio-kulturellen Hintergrund und höhere kognitive Grundfähigkeiten aufwiesen als Förderschülerinnen und -schüler. Auch bei statistischer Kontrolle dieser vergleichsweise günstigeren individuellen Lernvoraussetzungen scheinen jene jedoch von einer Beschulung in Regelschulen zu profitieren. Daraus könnte geschlussfolgert werden, dass eine Beschulung von Kindern mit SPF in Förderschulen mit der Verstärkung von sozialen Disparitäten einhergehen kann. Dies weist darauf hin, dass Kinder mit SPF, die in einer Förderschule unterrichtet werden, doppelt benachteiligt sind (vgl. Schümer 2004). Zum einen weisen Kinder mit SPF aus sozial schwachen Familien aufgrund ihrer individuellen Lernausgangsbedingungen niedrigere schulische Kompetenzen in Schulleistungstests auf. Zum anderen werden sie auch häufiger in Förderschulen unterrichtet, in denen Kinder mit SPF nach Kontrolle individueller Schülermerkmale niedrigere Kompetenzstände erreichen. Das kann letztlich dazu führen, dass sich herkunftsbedingte Bildungsnachteile institutionell verstärken (Schümer 2004, S. 102).

4.2.6.2 Erklärungsansätze für den Einfluss der Beschulungsart auf die schulischen Kompetenzen von Kindern mit SPF

Da der Fokus der Untersuchung auf der Quantifizierung von Effekten des Besuchs einer Regelschule auf die Schulleistung von Schülerinnen und Schülern mit SPF lag, lassen sich über die den Effekten zugrundeliegenden *Mechanismen* keine empirisch abgesicherten Aussagen treffen. Eine mögliche Ursache könnte darin liegen, dass Regelschulen im Vergleich zu Förderschulen günstigere Lern- und Entwicklungsmöglichkeiten für Kinder mit SPF bereitstellen, die unabhängig und zusätzlich zu individuellen Lernausgangsbedingungen schulische Leistungsentwicklungen begünstigen (Baumert et al. 2006; Bos et al. 2010; Wocken 2005). Diese differenzielle Leistungsentwicklung kann zum einen durch Effekte der sozialen und leistungsbezogenen Zusammensetzung der Schülerschaft verursacht werden (Stanat et al. 2010). Sowohl Auswertungen amtlicher Schulstatistiken als auch empirische Studien konnten zeigen, dass Kinder in Förderschulen (vor allem mit dem Förderschwerpunkt *Lernen*) häufiger aus sozial schwachen und bildungsfernen Familien stammen als Kinder in Regelschulen. Außerdem sind Kinder mit einem Zuwanderungshintergrund an Förderschulen überrepräsentiert (Dietze 2011; Opp et al. 2008; Wocken 2000, 2005). Unterschiede in der sozialen und kulturellen Zusammensetzung der Schülerschaft können bereits bestehende

Leistungsnachteile verstärken und sich als *institutioneller Matthäuseffekt* (Baumert et al. 2006) auf die individuelle Leistungsentwicklung der Schülerinnen und Schüler auswirken. In Schulen mit einem hohen Anteil an Kindern aus bildungsfernen Schichten kann ein eher ungünstiges Lernmilieu entstehen, in dem schulischer Erfolg und Leistungsstreben weniger wertgeschätzt werden (Eccles und Roeser 2011). Zum anderen könnten Kinder mit SPF durch Interaktionen mit leistungstärkeren Mitschülerinnen und Mitschülern in Regelschulen stärkere Lernfortschritte erzielen (Hanushek et al. 2003; Slavin 1996).

Dass die positiven Effekte des Besuchs einer Regelschule für Kinder mit dem Förderschwerpunkt *Lernen* offenbar deutlich größer sind als für Kinder mit dem Förderschwerpunkt *Sprache* könnte unter anderem auf Unterschiede zwischen den beiden Schülergruppen und in den Lehrplänen für die beiden Förderschwerpunkte zurückzuführen sein. Der Zuweisung zu den Förderschwerpunkten liegen unterschiedliche Diagnosekriterien zugrunde. Kinder mit dem Förderschwerpunkt *Lernen* weisen Beeinträchtigungen im schulischen Lernen und Leistungsverhalten auf, die durch vielfältige kognitive, sensorische aber auch soziale und emotionale Faktoren verursacht sein können (KMK 1999). Im Gegensatz dazu sind die Diagnosekriterien für eine Zuweisung zum Förderschwerpunkt *Sprache* enger umrissen und beziehen sich auf funktionale sprachliche Probleme, die sich sowohl in der Kommunikation als auch in altersunüblichen Leistungen im schriftlichen und sprachlichen Bereich äußern (KMK 1998). Möglicherweise verfügen Kinder mit dem Förderschwerpunkt *Sprache* daher über günstigere Lernausgangsbedingungen als Kinder mit dem Förderschwerpunkt *Lernen* (s. auch Hanushek et al. 2002). Tatsächlich ergaben zusätzliche Analysen unserer Daten, dass die Kinder mit dem Förderschwerpunkt *Sprache* über höhere kognitive Grundfähigkeiten verfügten als die Kinder mit dem Förderschwerpunkt *Lernen* ($d = 0,99$).

Weiterhin orientieren sich die Lehrpläne an Förderschulen mit dem Förderschwerpunkt *Sprache* insbesondere im mathematischen Bereich stärker an denen der Regelschule als die Lehrpläne der Förderschulen mit dem Förderschwerpunkt *Lernen* (KMK 1998). Die Zusammenfassung leistungsschwacher Schülerinnen und Schüler in homogenen Lerngruppen führt insbesondere dann zu negativen Auswirkungen auf ihre Schulleistungen, wenn diese mit geringeren curricularen Erwartungen einhergeht (Hattie 2002; Schofield 2010).

4.2.6.3 Einschränkungen der Untersuchung und Forschungsausblick

Eine Einschränkung der vorliegenden Untersuchung besteht darin, dass der Schweregrad der Beeinträchtigung der Kinder mit SPF nicht erhoben wurde. Dabei handelt es sich um ein generelles Problem bei der Untersuchung von Schülerinnen und Schülern mit SPF im Rahmen von Large-Scale-Assessments (Lindsay 2007; Ruijs und Peetsma 2009). Üblicherweise existieren lediglich globale Ausschlusskriterien für die Teilnahme an den Leistungstests. Am IQB-Ländervergleich 2011 nahmen nur Kinder teil, von denen angenommen werden konnte, dass sie die Kompetenztests bearbeiten können. Ausgeschlossen wurden Schülerinnen und Schüler mit einer körperlichen oder geistigen Beeinträchtigung sowie Kinder nichtdeutscher Muttersprache, die weniger als ein Jahr in deutscher Sprache unterrichtet worden sind. Außerdem wurden Kinder ausgeschlossen, die durch die Testsituation und die Testanweisungen massiv emotional belastet worden wären (Richter et al. 2012). Von diesen Kriterien lässt sich die Vermutung ableiten, dass die vorliegende Stichprobe überwiegend aus Kindern mit leichten bis moderaten Lern- bzw. Sprachbeeinträchtigungen bestand. Unklar ist allerdings, ob sich die beiden Untersuchungsgruppen im durchschnittlichen Schweregrad der Beeinträchtigung unterscheiden. Dies sollte in zukünftigen Studien berücksichtigt werden.

In der vorliegenden Untersuchung konnte zwar der Schweregrad des SPF nicht kontrolliert werden, Unterschiede in den kognitiven Grundfähigkeiten der Schülerinnen und Schüler wurden im Rahmen des Propensity Score Matchings jedoch berücksichtigt. Es ist allerdings möglich, dass auch die kognitiven Grundfähigkeiten der Schülerinnen und Schüler durch die Art der Beschulung beeinflusst worden sind – dies lässt sich aufgrund der querschnittlichen Anlage der Studie nicht ausschließen (Becker et al. 2012, Baumert et al. 2006). Zur empirischen Absicherung der berichteten Ergebnisse sind deshalb längsschnittliche Analysen zur Kompetenzentwicklung von Kindern mit SPF in Regel- und Förderschulen unerlässlich. Vor dem Hintergrund der Heterogenität der Schülergruppe mit SPF bedarf es in zukünftigen Studien zudem detaillierterer Analysen zur Frage, welche Schülerinnen und Schüler mit SPF durch Unterricht in einer Regelschule besonders gefördert werden können. So könnte untersucht werden, ob sich der gemeinsame Unterricht in Regelschulen je nach Schweregrad der Beeinträchtigung unterschiedlich auf die schulischen Kompetenzen von Kindern mit SPF auswirkt. Eine weitere offene Frage ist, inwieweit die Kompetenzentwicklung von Schülerinnen und Schülern *ohne* SPF durch integrativen bzw. inklusiven Unterricht beeinflusst wird. Die bislang durchgeführten empirischen Untersuchungen konnten für diese Schülergruppe überwiegend keine negativen Effekte identifizieren (Hanushek et al. 2002; Farrell et al. 2007;

Ruijs et al. 2010). Allerdings scheint der Zusammenhang zwischen dem Anteil von Kindern mit SPF in Regelschulen und der Schulleistung ihrer Mitschülerinnen und Mitschüler von weiteren Merkmalen der Schule und der Unterrichtsgestaltung, wie etwa personalen Ressourcen und Klassengröße, abhängig zu sein (Dyson et al. 2004; Farrell et al. 2007). Außerdem könnten die Auswirkungen auf die Leistungsentwicklung der Kinder *ohne* SPF auch je nach Förderschwerpunkt der integrativ bzw. inklusiv beschulten Kinder mit SPF variieren (Farrell et al. 2007).

Als Fazit lässt sich festgehalten, dass die vorliegende Untersuchung – trotz der dargestellten Einschränkungen – empirische Hinweise auf einen positiven Einfluss der Beschulung von Schülerinnen und Schülern mit SPF in Regelschulen liefert. Die Schätzung der Effekte erfolgte dabei anhand von Daten einer bundesweit durchgeführten Large-Scale-Assessment-Studie mit vergleichsweise großem Stichprobenumfang. Wichtiger Forschungsbedarf besteht jedoch weiterhin im Hinblick auf die Frage, wie ein gemeinsamer Unterricht von Kindern mit SPF in Regelschulen im Detail ausgestaltet werden sollte, um für alle Schülerinnen und Schüler effektiv zu sein (Zigmond 2003). Dies ist nicht zuletzt auch für evidenzgestützte Entscheidungen bei der weiteren Umsetzung inklusiver Bemühungen im deutschen Bildungssystem von besonderer Bedeutung.

Literaturverzeichnis

- Agirdag, Orhan, Mieke Van Houtte, und Piet Van Avermaet. 2012. Why does the ethnic and socio-economic composition of schools influence math achievement? The role of sense of futility and futility culture. *European Sociological Review* 28: 366–378.
- Baker, Edward T., Margaret C. Wang, und Herbert J. Walberg. 1994. The effects of inclusion on learning. *Educational Leadership* 52: 33–35.
- Barth, Joan M., Sarah T. Dunlap, Heather Dane, John E. Lochman, und Karen C. Wells. 2004. Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology* 42: 115–133.
- Baumert, Jürgen et al. 2008. *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, Jürgen, Kai Maaz, Petra Stanat, und Rainer Watermann. 2009. Schulkomposition oder Institution - was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die deutsche Schule* 101: 33–46.
- Baumert, Jürgen, Peter M. Roeder, Fritz Sang, und Bernhard Schmitz. 1986. Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. *Zeitschrift für Pädagogik* 5: 639–660.
- Baumert, Jürgen, Petra Stanat, und Rainer Watermann. 2006. Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*, Hrsg. Jür-

- gen Baumert, Petra Stanat, und Rainer Watermann, 95–188. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bear, George G., Kathleen M. Minke, und Maureen A. Manning. 2002. Self-concept of students with learning disabilities: A meta-analysis. *School Psychology Review* 31: 405–27.
- Becker, Michael. 2011. Matching-Verfahren und Gruppenvergleiche. In *Enzyklopädie Erziehungswissenschaft Online (EEO), Fachgebiet Methoden der empirischen erziehungswissenschaftlichen Forschung*, Hrsg. Sabine Maschke und Ludwig Stecher. Weinheim: Juventa.
- Becker, Michael, Oliver Lüdtke, Ulrich Trautwein, Olaf Köller, und Jürgen Baumert. 2012. The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology* 104: 682–699.
- Behrens, Ulrike, Katrin Böhme, und Michael Krelle. 2009. Zuhören – Operationalisierung und fachdidaktische Implikationen. In *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*, Hrsg. Dietlinde Granzer et al., 357–375. Weinheim: Beltz.
- Biewer, Gottfried. 2010. *Grundlagen der Heilpädagogik und inklusiven Pädagogik*. Bad Heilbrunn: Klinkhardt.
- Böhme, Katrin, und Albert Bremerich-Vos. 2012. Die im Ländervergleich 2011 in den Fächern Deutsch und Mathematik untersuchten Kompetenzen: Beschreibung der im Fach Deutsch untersuchten Kompetenzen. In *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*, Hrsg. Petra Stanat, Hans Anand Pant, Katrin Böhme, und Dirk Richter, 19–33. Münster: Waxmann.
- Bos, Wilfried, Sabrina Müller, und Tobias C. Stubbe. 2010. Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In *Bildungsverlierer. Neue Ungleichheiten*, Hrsg. Gudrun Quenzel und Klaus Hurrelmann, 375–397. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boudon, Raymond. 1974. *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: John Wiley & Sons.
- Bremerich-Vos, Albert, und Katrin Böhme. 2009. Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*, Hrsg. Dietlinde Granzer et al., 219–249. Weinheim: Beltz.
- BRK. 2008. *Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention) vom 21. 12. 2008*. New York: Vereinte Nationen.
- Buuren van, Stef, und Karin Groothuis-Oudshoorn. 2011. MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software* 45: 1–68.
- Carlberg, Conrad, und Kenneth Kavale. 1980. The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *The Journal of Special Education* 14: 295–309.
- Cohen, Jacob. 1992. A power primer. *Psychological Bulletin* 112: 155–159.
- Cole, Cassandra M, Nancy Waldron, und Massoumeh Majd. 2004. Academic progress of students across inclusive and traditional settings. *Mental retardation* 42: 136–144.
- Dar, Yehezkel, und Nura Resh. 1986. Classroom intellectual composition and academic achievement. *American Educational Research Journal* 23: 357–374.
- Dietze, Torsten. 2011. Sonderpädagogische Förderung in Zahlen - Ergebnisse der Schulstatistik 2009/10 mit einem Schwerpunkt auf der Analyse regionaler Disparitäten. *Zeitschrift für Inklusion* 6: 1–21.
- Dietze, Torsten. 2012. Zum Stand der sonderpädagogischen Förderung in Deutschland. Die Schulstatistik 2010/11. *Zeitschrift für Heilpädagogik* 63: 26–31.
- Dyson, Alan, Peter Farrell, Filiz Polat, Graeme Huteson und Frances Gallannaugh. 2004. *Inclusion and pupil achievement*. London: Department for Education and Skills.
- Eccles, Jacquelynne S., und Robert W. Roeser. 2011. Schools as developmental contexts during adolescence. *Journal of Research on Adolescence* 21: 225–241.
- Farrell, Peter, Alan Dyson, Filiz Polat, Graeme Huteson, und Frances Gallannaugh. 2007. The relationship between inclusion and academic achievement in English mainstream schools. *School Effectiveness and School Improvement* 18: 335–352.

- Gamoran, Adam. 1986. Instructional and institutional effects of ability grouping. *Sociology of Education* 59: 185–198.
- Ganzeboom, Harry B.G., Paul M. De Graaf, und Donald J. Treiman. 1992. A standard international socio-economic index of occupational status. *Social Science Research* 21: 1–56.
- Gomolla, Mechthild. 2006. Fördern und Fordern allein genügt nicht! Mechanismen institutioneller Diskriminierung von Migrantenkindern und -jugendlichen im deutschen Schulsystem. In *Schief-lagen im Bildungssystem*, Hrsg. Georg Auernheimer, 87–102. VS Verlag für Sozialwissenschaften.
- Granzer, Dietlinde et al. 2009. *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*. Weinheim: Beltz.
- Haeberlin, Urs. 1991. Die Integration von leistungsschwachen Schülern. Ein Überblick über empirische Forschungsergebnisse zu Wirkungen von Regelklassen, Integrationsklassen und Sonderklassen auf „Lernbehinderte“. *Zeitschrift für Pädagogik* 37: 167–189.
- Hallinan, Maureen T., und Warren N. Kubitschek. 1999. Curriculum differentiation and high school achievement. *Social Psychology of Education* 3: 41–62.
- Hansen, Ben B. 2004. Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association* 99: 609–618.
- Hanushek, Eric A., John F. Kain, Jacob M. Markman, und Steven G. Rivkin. 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics* 18: 527–544.
- Hanushek, Eric A., John F. Kain, und Steven G. Rivkin. 2002. Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics* 84: 584–599.
- Harrington, Stuart A. 1997. Full inclusion for students with learning disabilities: A review of the evidence. *School Community Journal* 7: 63–71.
- Hattie, John A.C. 2002. Classroom composition and peer effects. *International Journal of Educational Research* 37: 449–481.
- Heller, Kurt A., und Christoph Perleth. 2000. *KFT 4-12+R: kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test.
- Hinz, Andreas. 2002. Von der Integration zur Inklusion - terminologisches Spiel oder konzeptionelle Weiterentwicklung? *Zeitschrift für Heilpädagogik* 53: 354–361.
- Ho, Daniel E., Kosuke Imai, Gary King, und Elizabeth A. Stuart. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15: 199–236.
- Ho, Daniel E., Kosuke Imai, Gary King, und Elizabeth A. Stuart. 2011. MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software* 42: 1–28.
- Hocutt, Anne M. 1996. Effectiveness of special education: Is placement the critical factor? *The Future of Children* 6: 77–102.
- Hornstra, Lisette, Eddie Denessen, Joep Bakker, Linda van den Bergh, und Marinus Voeten. 2010. Teacher attitudes toward dyslexia: Effects on teacher expectations and the academic achievement of students with dyslexia. *Journal of Learning Disabilities* 43: 515–529.
- Imai, Kosuke, Gary King, und Olivia Lau. 2008. Toward a common framework for statistical analysis and development. *Journal of Computational Graphics and Statistics* 17: 1–22.
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. 1999. *Empfehlungen zum Förderschwerpunkt Lernen. Beschluss vom 01.10.1999*. <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopale.pdf> (Zugegriffen Mai 31, 2013).
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. 1998. *Empfehlungen zum Förderschwerpunkt Sprache. Beschluss vom 26.06.1998*. <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sprache.pdf> (Zugegriffen Mai 31, 2013).
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. 1994. *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluß vom 06.05.1994*. <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopae94.pdf> (Zugegriffen Mai 31, 2013).

- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. 2010. Pädagogische und rechtliche Aspekte der Umsetzung des Übereinkommens der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (Behindertenrechtskonvention - VN-BRK) in der schulischen Bildung. Beschluss vom 18.11.2010.
http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_11_18-Behindertenrechtskonvention.pdf (Zugegriffen Mai 31, 2013).
- Kuhl, Poldi, Sebastian Weirich, Nicole Haag, und Aleksander Kocaj. 2013. Zur Validität und Messinvarianz bei der Erfassung der Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Large-Scale-Assessments. Kiel: Vortrag auf der ersten Tagung der Gesellschaft für empirische Bildungsforschung (GEBF).
- Lehmann, Rainer H. 2006. Zur Bedeutung der kognitiven Heterogenität von Schulklassen für den Lernstand am Ende der Klassenstufe 4. In *Risikofaktoren kindlicher Entwicklung*, Hrsg. Agi Schröder-Lenzen, 109–121. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lehmann, Rainer, und Ellen Hoffmann, Hrsg. 2009. *BELLA: Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf „Lernen“*. Münster: Waxmann.
- Lentz, Astrid, und Frank-Olaf Radtke. 1994. Bildungsghettos - Institutionalisierte Diskriminierung von Migrantenkindern in der Grundschule. *Unterrichtswissenschaft* 22: 182–191.
- Lindsay, Geoff. 2007. Educational psychology and the effectiveness of inclusive education/mainstreaming. *British Journal of Educational Psychology* 77: 1–24.
- Maaz, Kai, Ulrich Trautwein, Oliver Lüdtke, und Jürgen Baumert. 2008. Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives* 2: 99–106.
- Maaz, Kai, Jürgen Baumert, und Ulrich Trautwein. 2010. Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In *Bildungsentscheidungen*, Hrsg. Jürgen Baumert, Kai Maaz, und Ulrich Trautwein, 11–46. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Markussen, Eifred. 2004. Special education: Does it help? A study of special education in Norwegian upper secondary schools. *European Journal of Special Needs Education* 19: 33–48.
- Morgan, Paul L., Michelle L. Frisco, George Farkas, und Jacob Hibel. 2010. A propensity score matching analysis of the effects of special education services. *The Journal of Special Education* 43: 236–254.
- Morgan, Stephen L., und Christopher Winship. 2007. *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge: Cambridge University Press.
- Myklebust, Jon Olav. 2006. Class placement and competence attainment among students with special educational needs. *British Journal of Special Education* 33: 76–81.
- Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78: 691–692.
- OECD. 1999. *Classifying educational programmes. Manual for ISCED-97 implementation in OECD countries*. Paris: Organisation for Economic Co-operation and Development.
- Opp, Günther, Ines Budnik, und Michael Fingerle. 2008. Sonderschulen — integrative Beschulung. In *Handbuch der Schulforschung*, Hrsg. Werner Helsper und Jeanette Böhme, 341–361. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Peetsma, Thea, Margaretha Vergeer, Jaap Roeleveld, und Sjoerd Karsten. 2001. Inclusion in education: Comparing pupils' development in special and regular education. *Educational Review* 53: 125–135.
- Pfahl, Lisa. 2012. Bildung, Behinderung und Agency. Eine wissenssoziologische Untersuchung der Folgen schulischer Segregation und Inklusion. In *Soziologische Bildungsforschung, Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderhefte*, Hrsg. Rolf Becker und Heike Solga, 415–436. Wiesbaden: Springer.
- Preuss-Lausitz, Ulf. 2001. Gemeinsamer Unterricht Behinderter und Nichtbehinderter. *Zeitschrift für Erziehungswissenschaft* 4: 209–224.
- R Development Core Team. 2010. *R: A language and environment for statistical computing*. Wien: R Foundation for Statistical Computing <http://www.R-project.org>.

- Reiss, Kristina, und Henrik Winkelmann. 2009. Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*, Hrsg. Dietlinde Granzer et al., 120–141. Weinheim: Beltz.
- Richter, Dirk et al. 2012. Anlage und Durchführung des Ländervergleichs. In *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*, Hrsg. Petra Stanat, Hans Anand Pant, Katrin Böhme, und Dirk Richter, 85–102. Münster: Waxmann.
- Rogers, Donna Power, und Iris Marie Thiery. 2003. Does an inclusive setting affect reading comprehension in students with learning disabilities? Biloxi, MS
<http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED482684> (Zugegriffen Mai 31, 2013).
- Roppelt, Alexander, und Kristina Reiss. 2012. Die im Ländervergleich 2011 in den Fächern Deutsch und Mathematik untersuchten Kompetenzen: Beschreibung der im Fach Mathematik untersuchten Kompetenzen. In *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*, Hrsg. Petra Stanat, Hans Anand Pant, Katrin Böhme, und Dirk Richter, 34–44. Münster: Waxmann.
- Rosenbaum, Paul R. 1991. A characterization of optimal designs for observational studies. *Journal of the Royal Statistical Society. Series B (Methodological)* 53: 597–610.
- Rosenbaum, Paul R., und Donald B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39: 33–38.
- Rosenbaum, Paul R., und Donald B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.
- Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688–701.
- Rubin, Donald B. 1987. *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- Ruijs, Nienke M., und Thea T.D. Peetsma. 2009. Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review* 4: 67–79.
- Ruijs, Nienke, Thea Peetsma, und Ineke van der Veen. 2010. The presence of several students with special educational needs in inclusive education and the functioning of students with special educational needs. *Educational Review* 62: 1–37.
- Schafer, Joseph L. 1997. *Analysis of incomplete multivariate data*. 1. Aufl. London: Chapman & Hall.
- Schafer, Joseph L., und Joseph Kang. 2008. Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods* 13: 279–313.
- Scharenberg, Katja. 2012. *Leistungsheterogenität und Kompetenzentwicklung: Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Hrsg. Rolf Becker et al. Münster: Waxmann.
- Schofield, Janet W. 2010. International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. *Teachers College Record* 112: 1492–1528.
- Schor, Bruno J. 2003. Die Differenziertheit des deutschen Bildungswesens. Chance oder Hemmnis. *Zeitschrift für Heilpädagogik* 54: 370–376.
- Schumann, Brigitte. 2007. *"Ich schäme mich ja so!" Die Sonderschule für Lernbehinderte als "Schonraumfalle"*. Bad Heilbrunn: Klinkhardt.
- Schümer, Gundel. 2004. Zur doppelten Benachteiligung von Schülern aus unterprivilegierten Gesellschaftsschichten im deutschen Schulwesen. In *Die Institution Schule und die Lebenswelt der Schüler. Vertiefende Analysen der PISA-2000-Daten zum Kontext von Schülerleistungen*, Hrsg. Gundel Schümer, Klaus-Jürgen Tillmann, und Manfred Weiss, 73–114. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Slavin, Robert E. 1990. Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research* 60: 471–499.
- Slavin, Robert E. 1996. Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology* 21: 43–69.

- Stanat, Petra, Hans Anand Pant, Katrin Böhme, und Dirk Richter, Hrsg. 2012. *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stanat, Petra, Knut Schwippert, und Carola Gröhlich. 2010. Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb. Längsschnittliche Überprüfung eines umstrittenen Effekts. In *Migration, Identität, Sprache und Bildungserfolg*. 55. Beiheft der Zeitschrift für Pädagogik, Hrsg. Cristina Allemann-Ghionda, Petra Stanat, Kerstin Göbel, und Charlotte Röhner, 147–164. Weinheim: Beltz.
- Stuart, Elizabeth A. 2010. Matching methods for causal inference: A review and a look forward. *Statistical Science* 25: 1–21.
- Sullivan, Amanda L., und Samuel Field. 2013. Do preschool special education services make a difference in kindergarten reading and mathematics skills?: A propensity score weighting analysis. *Journal of School Psychology* 51: 243–260.
- Tent, Lothar, Matthias Witt, Christiane Zschoche-Lieberum, und Wolfgang Bürger. 1991. Über die pädagogische Wirksamkeit der Schule für Lernbehinderte. *Zeitschrift für Heilpädagogik* 5: 289–320.
- Vaughn, Sharon, Batya E. Elbaum, und Jeanne Shay Schumm. 1996. The effects of inclusion on the social functioning of students with learning disabilities. *Journal of Learning Disabilities* 29: 599–608.
- Walberg, Herbert J., und Shiow-Ling Tsai. 1983. Matthew effects in education. *American Educational Research Journal* 20: 359–373.
- Wang, Margaret C., und Edward T. Baker. 1985. Mainstreaming programs: Design features and effects. *The Journal of Special Education* 19: 503–521.
- Wocken, Hans. 2005. *Andere Länder, andere Schüler? Vergleichende Untersuchung von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen*. <http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf> (Zugegriffen Mai 31, 2013).
- Wocken, Hans. 2000. Leistung, Intelligenz und Soziallage von Schülern mit Lernbehinderungen. Vergleichende Untersuchungen an Förderschulen in Hamburg. *Zeitschrift für Heilpädagogik* 51: 492–503.
- Wocken, Hans, und Carola Gröhlich. 2009. Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen. In *KESS 7 - Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*, Hrsg. Wilfried Bos, Martin Bensen, und Carola Gröhlich, 133–142. Münster: Waxmann.
- Zigmond, Naomi. 2003. Where should students with disabilities receive special education services? Is one place better than another? *The Journal of Special Education* 37: 193–199.

4.3 Teilstudie 3: Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf

Abstract

Ein Ziel aktueller Inklusionsbemühungen ist es, Bildungsbenachteiligungen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) abzubauen, die mit ihrer Beschulung in Förderschulen und mit ihrem sozio-kulturellen Hintergrund verbunden sind. In der vorliegenden Studie wurde anhand von Daten des *IQB-Ländervergleichs Primarstufe 2011* geprüft, ob Merkmale der Klassenkomposition zu Unterschieden in den schulischen Kompetenzen (Mathematik, Leseverständnis) von Kindern mit SPF in Förderschulen ($N = 413$) und in sonstigen allgemeinbildenden Grundschulen ($N = 658$) beitragen. Signifikante Zusammenhänge zwischen den kognitiven Grundfähigkeiten auf Klassenebene und den schulischen Kompetenzen auf Individualebene nach Kontrolle der Beschulungsart weisen darauf hin, dass Kinder mit SPF in Klassen mit insgesamt höherem Fähigkeitsniveau bessere schulische Kompetenzen erzielen. Die Ergebnisse geben Hinweise darauf, dass allgemeine Grundschulen im Vergleich zu Förderschulen für Kinder mit SPF bessere Lern- und Entwicklungsmöglichkeiten bereitstellen könnten.

Schlagwörter: Inklusion, sonderpädagogischer Förderbedarf, Klassenkomposition, schulische Kompetenzen, Primarstufe

Teilstudie 3 ist in folgender Publikation erschienen:

Kocaj, A., Kuhl, P., Rjosk, C., Jansen, M., Pant, H. A. & Stanat, P. (2015). Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf. In: P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 335–370). Wiesbaden: Springer VS.

4.3.1 Einleitung

Ein wichtiges Ziel inklusiver Bemühungen im deutschen Bildungssystem besteht darin, die Lernumwelt von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) so zu gestalten, dass Voraussetzungen für eine „bestmögliche schulische und soziale Entwicklung“ geschaffen werden (Behindertenrechtskonvention der Vereinten Nationen, VN-BRK 2008, S. 37). Für die schulischen Kompetenzen, die einen wichtigen Aspekt der schulischen Entwicklung bilden, zeigen sich mehrheitlich Leistungsvorsprünge zugunsten der Schülerinnen und Schüler mit SPF in allgemeinen Schulen im Vergleich zu Kindern und Jugendlichen in Förderschulen (Ruijs & Peetsma 2009). So konnte kürzlich in einer bundesweiten Studie (Kocaj et al. 2014) festgestellt werden, dass – unter Berücksichtigung der differenziellen Eingangsselektivität der Schularten – Schülerinnen und Schüler mit SPF in sonstigen allgemeinbildenden Grundschulen (im Folgenden „Grundschulen“ genannt) höhere schulische Kompetenzen im Lesen, Zuhören und in Mathematik erzielen als vergleichbare Kinder mit SPF in Förderschulen.

Die Leistungsunterschiede zwischen Schülerinnen und Schülern mit SPF in Abhängigkeit von ihrer Beschulungsart können auf verschiedene Merkmale ihrer Lernumwelt zurückzuführen sein. Kinder mit SPF werden u.a. aufgrund ihrer schulischen Leistungsfähigkeit zu unterschiedlichen Schularten zugewiesen (Opp et al. 2008, Preuss-Lausitz 2001). Damit verbunden sind pädagogische Annahmen, dass die Unterrichtsgestaltung in leistungshomogenen Lerngruppen besser auf die Bedürfnisse der Schülerinnen und Schüler abgestimmt werden kann und ihre Lernentwicklung begünstigt (Köller et al. 2013, Preuss-Lausitz 2013). Durch die Zuweisung von Kindern mit SPF zu unterschiedlichen Schularten kommt es aber nicht nur zu einer leistungs- bzw. fähigkeitsbezogenen Trennung, sondern auch zu einer sozio-kulturellen Trennung der Schülerschaft (Klein 2001, Wocken 2005). Diese schulartspezifischen Unterschiede in der sozio-kulturellen und fähigkeitsbezogenen Zusammensetzung der Schülerschaft werden wesentlich dadurch bestimmt, dass Kinder mit SPF in Grundschulen gemeinsam mit Kindern ohne SPF unterrichtet werden. Somit können nicht nur Merkmale der Unterrichtsorganisation (u.a. Lehrpläne, Stundentafeln) in Abhängigkeit der Beschulungsart, sondern auch die fähigkeitsbezogene und sozio-kulturelle Zusammensetzung der Schulklasse, die ein Kind besucht, zu seiner schulischen Kompetenzentwicklung beitragen (Baumert et al. 2006). Es wird angenommen, dass Lehrkräfte ihre Unterrichtsmethoden, Erwartungen und Anforderungen an die Zusammensetzung der Schülerschaft in ihrer Klasse anpassen. Dem-

nach gestalten Lehrkräfte in Schulklassen mit einem hohen Anteil an leistungsschwachen Kindern ihren Unterricht weniger anspruchsvoll und verringern ihre Leistungserwartungen (Köller et al. 2013). Darüber hinaus kann sich die Zusammensetzung der Klasse auch über Interaktionsprozesse zwischen Schülerinnen und Schülern auf deren schulische Kompetenzen auswirken. Das kann zum einen durch geteilte Leistungsnormen und Lerneinstellungen und zum anderen durch soziale Lernprozessen vermittelt werden (Justice et al. 2014, Thrupp et al. 2002). Durch die Verteilung der Schülerinnen und Schüler mit SPF auf unterschiedliche Schularten können schulartspezifische Lernumwelten entstehen, die u.a. durch die fähigkeitsbezogene und sozio-kulturelle Zusammensetzung der Schulklassen geprägt sind und die schulische Leistungsentwicklung beeinflussen.

Empirische Anhaltspunkte dafür, ob auch Merkmale der Klassenkomposition zu Leistungsunterschieden zwischen Schülerinnen und Schülern mit SPF in Grund- und Förderschulen beitragen, sind bislang rar. An dieser Forschungsfrage setzt die vorliegende Untersuchung an. Es soll geprüft werden, ob nach Berücksichtigung der Klassenkomposition noch Effekte der Beschulungsart auf die schulischen Kompetenzen von Kindern mit SPF nachgewiesen werden können. Zunächst werden die Ergebnisse des Vergleichs zwischen Kindern mit SPF in Förder- und Grundschulen hinsichtlich demografischer Merkmale und sozio-kultureller Herkunft noch einmal dargestellt und um weitere Merkmale der sozio-kulturellen Herkunft ergänzt (vgl. Kocaj et al. 2014). Darauf aufbauend soll geprüft werden, ob Förderschulen im Vergleich zu Grundschulen differenzielle Lern- und Entwicklungsmilieus darstellen (Köller & Baumert 2001), in denen Kinder mit SPF in Förderschulen aufgrund kompositionsbezogener Effekte niedrigere schulische Kompetenzen aufweisen als vergleichbare Kinder mit SPF in Grundschulen.

4.3.2 Theoretischer Hintergrund

Förderschulen und sonstige allgemeinbildende Schulen unterscheiden sich primär in der leistungs- bzw. fähigkeitsbezogenen Zusammensetzung der Schülerschaft (Bos et al. 2010, Wocken 2005). Verschiedene Untersuchungen haben jedoch auch gezeigt, dass der Anteil von Kindern aus bildungsfernen und sozial schwachen Familien in Förderschulen besonders hoch ist (Autorengruppe Bildungsberichterstattung 2010, Klein 2001, Wocken 2005). Im ersten Teil werden zunächst Erklärungsansätze für dieses Muster dargestellt. Im zweiten Teil wird

darauf eingegangen, wie sich Unterschiede in der sozio-kulturellen und fähigkeitsbezogenen Klassenkomposition auf die Lernentwicklung von Kindern in Förderschulen im Vergleich zu Kindern mit SPF in Grundschulen auswirken können.

4.3.2.1 Koppelung der Förderschulzuweisung an den sozialen Hintergrund

Als mögliche Erklärungen für die Überrepräsentation von Kindern aus sozial schwachen Familien in Förderschulen lassen sich u. a. individuumszentrierte und interaktionale Ansätze heranziehen (Kornmann 2006, Kronig 2003).

Individuumszentrierte Ansätze beziehen sich auf die Lernvoraussetzungen und Lebenslagen von Kindern aus sozial schwachen Familien. Boudon (1974) unterscheidet in seinem Modell zur Entstehung sozialer Ungleichheiten im Bildungssystem primäre und sekundäre Effekte der sozialen Herkunft (vgl. Baumert & Schümer 2001, Maaz et al. 2010). Primäre Herkunftseffekte zeigen sich im Zusammenhang zwischen dem sozio-ökonomischem und kulturellen Kapital der Familien mit den schulischen Kompetenzen der Schülerinnen und Schüler. Leistungsunterschiede in Abhängigkeit von der sozialen Herkunft beruhen darauf, dass Kindern aus sozial besser gestellten Familien mehr Ressourcen und Lernanregungen im Elternhaus zur Verfügung stehen. So könnten Kinder aus sozial schwachen Familien u.a. deshalb geringere Schulleistungen aufweisen, weil das häusliche Anregungsniveau geringer ist als in sozial besser gestellten Familien (z. B. McElvany et al. 2009). Schulische Minderleistungen sind wiederum ein Ausgangspunkt für die Feststellung eines SPF und die Überweisung auf Förderschulen (Bos et al. 2010, Opp et al. 2008, Preuss-Lausitz 2001). Dadurch können primäre Herkunftseffekte zur Überrepräsentation von Kindern aus sozial schwachen und bildungsfernen Familien in Förderschulen beitragen.

Sekundäre Herkunftseffekte bestehen nach Boudon (1974), wenn Kinder aus sozial schwachen Familien trotz vergleichbarer Schulleistungen niedrigere Bildungsabschlüsse erzielen als Kinder aus sozial besser gestellten Familien. Solche sozialen Herkunftseffekte können u. a. durch risikoscheuere Kosten-Nutzen-Bewertungen und durch geringere elterliche Bildungsaspirationen für die Kinder entstehen (Stocké 2007). Sekundäre Effekte spiegeln somit stärker spezifische Bildungsentscheidungen der Eltern wider. Sie zeigen sich in der Art der Bewertung von Bildungsabschlüssen und im Kenntnisstand über Aufstiegsmöglichkeiten im Bildungswesen, die je nach sozio-kulturellem Hintergrund variieren (Maaz et al. 2010).

Beispielsweise könnten Eltern aus unteren sozialen Schichten bei einer in Frage stehenden Zuweisung ihres Kindes auf eine Förderschule ihre Handlungsspielräume schlechter kennen und zudem die Reichweite der Konsequenzen einer solchen Zuweisung für die zukünftigen Bildungschancen des Kindes unterschätzen.

Interaktionale Erklärungsansätze gehen davon aus, dass die Überweisung von Schülerinnen und Schülern aus sozial schwachen Familien auf Förderschulen durch soziale Stereotype und Erwartungshaltungen der Lehrkräfte beeinflusst wird (Diamond et al. 2004, Rist 2000, Tenenbaum & Ruck 2007). Ausgangspunkt dieser Annahme ist, dass die eingesetzten Diagnoseverfahren zur Feststellung eines SPF bzw. zur Überweisungsentscheidung mit einer gewissen Unsicherheit verbunden sind. Das trifft vor allem auf den Förderschwerpunkt *Lernen* zu (Kornmann 2006, Kronig 2001, 2003). Dadurch fließen in Übergangsentscheidungen neben diagnostischen Testergebnissen auch subjektiven Einschätzungen wie z. B. Schulnoten und Urteile über das Sozialverhalten und die Gesamtentwicklung des betroffenen Kindes ein (Lentz & Radtke 1994). Diese Einschätzungen könnten bei Kindern aus sozial besser gestellten Familien günstiger ausfallen, was zur Konsequenz hätte, dass sie bei ansonsten ähnlichen Lernvoraussetzungen mit einer geringeren Wahrscheinlichkeit auf eine Förderschule überwiesen werden als Kinder aus sozial schwachen Familien (Gomolla 2006, Maaz et al. 2010).

4.3.2.2 Entstehung schulartspezifischer Lern- und Entwicklungsmilieus

Unterschiedliche Leistungsentwicklungen in Abhängigkeit von der besuchten Schulart können auf mindestens zwei Wirkmechanismen zurückgehen (Baumert et al. 2006, Dumont et al. 2013, Harris 2010, Köller & Baumert 2001). Erstens wird angenommen, dass Unterschiede in der Unterrichtsorganisation (z. B. in Stundentafeln und Curricula) und in der Unterrichtsgestaltung (z. B. didaktische Traditionen, Fachwissen der Lehrkräfte) zwischen Förder- und Grundschulen eine differenzielle Leistungsentwicklung bedingen können (*institutionelle Effekte*, vgl. Köller & Baumert 2001, Schofield 2010). Wocken (2005) geht davon aus, dass die Lehrpläne in Förderschulen (mit dem Förderschwerpunkt *Lernen*) weniger umfangreich und anspruchsvoll sind als in sonstigen allgemeinbildenden Schulen. Die stärkere Ausrichtung des Unterrichts auf den akademischen Fortschritt in Grundschulen im Vergleich zu Förderschulen sollte sich positiv auf die schulischen Leistungen der Kinder mit SPF auswirken (Markussen 2004, Ruijs & Peetsma 2009).

Zweitens kommt auch der Klassenkomposition eine wichtige Bedeutung für Leistungsunterschiede zwischen Kindern mit SPF in Förder- bzw. Grundschulen zu (Harris 2010, Thrupp et al. 2002, Rauer & Schuck 2007). Dabei sind Effekte der Klassenkomposition nach Köller und Baumert (2001) mit institutionellen Effekten der Schulart verknüpft. Schulartspezifische Lernmilieus können zum einen dadurch entstehen, dass Schülerinnen und Schüler auf Grundlage ihrer schulischen Leistungsfähigkeit zu unterschiedlichen Schularten zugewiesen werden. Zum anderen kann auch die Kopplung der Schulartzuweisungen an den sozialen Hintergrund der Eltern zur Entstehung spezifischer Lernmilieus in Förder- und Grundschulen beitragen (Rauer & Schuck 2007, Wocken 2005). Diese spiegeln sich sowohl in der leistungs- bzw. fähigkeitsbezogenen als auch in der sozio-kulturellen Klassenkomposition wider.

Effekte der Klassenkomposition manifestieren sich auf der einen Seite in differenziellen Interaktionsmustern zwischen den Schülerinnen und Schülern und ihren Lehrkräften und äußern sich in der Anpassung des Anforderungsniveaus (Harris 2010, Weinert & Helmke 1997) und in der Anpassung der Erwartungshaltungen der Lehrkräfte (Diamond et al. 2004, Kristen 2002, Rist 2000), sowie dem Lernklima im Klassenzimmer (Baumert et al. 2009). Mehrere Ansätze gehen davon aus, dass Lehrkräfte in Klassen mit einem hohen Anteil an Schülerinnen und Schülern aus sozial schwachen Familien niedrigere Erwartungen hinsichtlich deren Leistungsvermögen entwickeln und in leistungsschwächeren Klassen das Anforderungsniveau absenken (Diamond et al. 2004, Gamoran 1986, Hornstra et al. 2010, Köller et al. 2013). Das kann mit einem geringeren Lernfortschritt der Schülerinnen und Schüler einhergehen (Diamond et al. 2004, Rist 2000).

Auf der anderen Seite wird davon ausgegangen, dass auch Interaktionen zwischen Schülerinnen und Schülern einer Klasse einen bedeutsamen Einfluss auf die schulischen Leistungen ausüben (*Peer Effekte*, Thrupp et al. 2002). Zum einen kann durch die Häufung von Kindern aus bildungsnahen Schichten ein günstiges Lernklima entstehen, in dem schulischer Erfolg und Leistungsstreben stärker wertgeschätzt werden (Eccles & Roeser 2011, Kristen 2002). Zum anderen können leistungsschwächere Schülerinnen und Schüler vom gemeinsamen Unterricht mit leistungsstarken Kindern profitieren, indem sie Lösungsstrategien leistungsstarker Mitschülerinnen und Mitschüler übernehmen (Gamoran 1986, Hallinan & Kubitschek 1999, Kristen 2002, Slavin 1996).

4.3.3 Empirischer Forschungsstand zu Auswirkungen differenzieller Lern- und Entwicklungsmilieus auf schulische Kompetenzen von Kindern mit SPF

Sowohl Auswertungen amtlicher Schulstatistiken als auch Befunde empirischer Studien lassen den Schluss zu, dass es sich bei Förderschülerinnen und -schülern um eine sozio-kulturell benachteiligte Gruppe handelt (Autorengruppe Bildungsberichterstattung 2010, Kocaj et al. 2014, Klein 2001, Wocken 2005). Außerdem weisen Schülerinnen und Schüler mit SPF in Grundschulen bessere Lernvoraussetzungen (z. B. kognitive Fähigkeiten) auf als Kinder in Förderschulen (Kocaj et al. 2014). Wie sich diese Unterschiede in Form von fähigkeitsbezogenen und sozio-kulturellen Kompositionseffekten auf die schulischen Kompetenzen von Kindern mit SPF auswirken, wurde im deutschsprachigen Raum bislang nur selten untersucht (Bos et al. 2010, Rauer & Schuck 2007, Wocken 2005).

Die Mehrzahl der Untersuchungen zu Effekten von Kompositionsmerkmalen auf individuelle schulische Kompetenzen bezieht sich auf Schülerinnen und Schüler ohne SPF. Diese Studien zeigen, dass insbesondere leistungsschwächere Schülerinnen und Schüler durch eine sozial ungünstige Klassenkomposition in ihren schulischen Leistungen beeinträchtigt werden (De Fraine et al. 2003, Rjosk et al. 2014). Den stärksten Einfluss auf die individuellen Leistungen in mathematischen und sprachlichen Bereichen übt jedoch das mittlere Leistungs- bzw. Fähigkeitsniveau der Klasse aus (Gröhlich et al. 2010, Nikolova 2011). Die Leistungsdifferenzierung durch die Aufteilung der Schülerinnen und Schüler in verschiedene institutionell getrennte Bildungsgänge geht auch mit einer sozialen Trennung der Schülerschaft einher (vgl. Baumert et al. 2006). So zeigte sich bei Baumert und Kollegen (2006) ein auf Schulebene hoher Zusammenhang ($R^2 = .70$) zwischen den durchschnittlichen Indexwerten der sozialen Herkunft (soziale Komposition) der Schülerinnen und Schüler und ihren mittleren kognitiven Grundfähigkeiten (fähigkeitsbezogene Komposition). In mathematischen und sprachlichen Kompetenzbereichen scheinen neben Merkmalen der Schülerzusammensetzung institutionelle Effekte eine Rolle für die Lernentwicklung von Schülerinnen und Schülern zu spielen.

Weiterhin verweisen Dumont und Kollegen (2013) auf den starken Zusammenhang zwischen Kompositionseffekten und institutionellen Effekten der Schulart im deutschen Bildungssystem. So fanden Köller und Baumert (2001), dass die Effekte des mittleren Leistungsniveaus im Sekundarbereich vollständig mit der Schulart konfundiert sind. Allerdings konnten Baumert und Kollegen (2006) auf Grundlage der PISA-Studie im Sekundarbereich zeigen, dass die Leistungskomposition auch nach Kontrolle für Schulart und soziale Zusammenset-

zung positiv mit der individuellen Schulleistung zusammenhängt. Es wird angenommen, dass sich diese Befunde auch auf die Förderschule übertragen lassen, da sich die soziale und leistungs- bzw. fähigkeitsbezogene Schülerzusammensetzung in Förderschulen ebenfalls deutlich von anderen Schularten unterscheidet (Bos et al. 2010). Eine gemeinsame Betrachtung von kompositionsbezogenen und institutionellen Effekten auf die schulischen Kompetenzen von Kindern mit SPF in Förder- und Grundschulen wurde bisher jedoch noch nicht vorgenommen.

Eine der wenigen relevanten Studien mit Schülerinnen und Schülern mit SPF ist die Untersuchung von Rauer und Schuck (2007), die anhand von Daten der KESS-4-Studie verschiedene Formen der integrativen Beschulung an Hamburger Grundschulen verglich. Die Autoren stellten fest, dass sich Leistungsunterschiede (in Deutsch, Mathematik, Naturwissenschaften, Englisch) bei Gruppen von Kindern mit SPF, die verschiedene Arten integrativen Unterrichts erhielten, teilweise auf Unterschiede im sozio-kulturellen Hintergrund der Schülerschaft zurückführen ließen (Rauer & Schuck 2007, vgl. auch Justice et al. 2014). In integrativ ausgerichteten Klassen mit einer günstigen sozio-kulturellen Schülerzusammensetzung waren die durchschnittlichen schulischen Leistungen von Kindern mit SPF besser als in integrativen Schulklassen mit einem hohen Anteil an Kindern aus bildungsfernen und sozial schwachen Familien (Rauer & Schuck 2007). Eine Limitation der Studie besteht allerdings darin, dass die Effekte der sozio-kulturellen Klassenzusammensetzung nur zwischen verschiedenen integrativen Unterrichtsformen und nicht-integrativen Grundschulklassen verglichen wurden, ohne Förderschulen einzubeziehen.

4.3.4 Fragestellung und Hypothesen

Aus dem dargestellten Forschungsstand ergeben sich zwei zentrale Forschungsfragen. Erstens soll untersucht werden, ob und ggf. wie sehr sich Schülerinnen und Schüler mit einem SPF in Förder- und Grundschulen in demografischen und sozio-kulturellen Merkmalen voneinander unterscheiden.¹¹ Zweitens soll der Frage nachgegangen werden, ob Förder- und Grundschulen differenzielle Lern- und Entwicklungsmilieus darstellen, in denen die schulischen Kompetenzen von Kindern mit SPF über ihre individuellen Lernausgangsbedingungen hinaus durch die

¹¹ Die Analysen basieren auf einer vorherigen Untersuchung (Kocaj et al. 2014) und wurden um die Merkmale der Lesezeit und beruflichen Situation der Eltern erweitert.

Klassenkomposition beeinflusst werden. In Einklang mit dem Forschungsstand sollte die fähigkeitsbezogene Klassenzusammensetzung einen stärkeren Einfluss auf die individuelle Schulleistung von Kindern mit SPF ausüben als die soziale Zusammensetzung der Schulklassen. Weiterhin soll geprüft werden, ob institutionelle Effekte der Beschulungsart auch noch mit den schulischen Kompetenzen der Kinder mit SPF zusammenhängen, wenn kompositionsbezogene Merkmale berücksichtigt werden. Dabei wird angenommen, dass sowohl institutionelle Effekte der Beschulungsart als auch Kompositionseffekte auftreten.

4.3.5 Methoden

Der Fragestellung zur Bedeutung der Klassenzusammensetzung auf die schulischen Kompetenzen von Schülerinnen und Schülern mit SPF wurde mit Hilfe von Mehrebenenanalysen nachgegangen. Dabei wurde geprüft, ob die Beschulungsart (*Institutionseffekt*) nach Berücksichtigung der fähigkeitsbezogenen und sozialen Klassenzusammensetzung (*Kompositionseffekte*) sowie individueller Schülermerkmale mit Unterschieden in den schulischen Kompetenzen von Kindern mit SPF zusammenhängen. Die schulischen Kompetenzen in Mathematik und im Leseverständnis stellen die abhängigen Variablen der Untersuchung dar. Die Beschulungsart (Förder- und Grundschulen) sowie die fähigkeitsbezogene und soziale Klassenzusammensetzung wurden als unabhängige Variablen in die Analysen aufgenommen.

Die vorliegenden Analysen basieren auf einer querschnittlichen Erhebung. Mögliche Leistungsunterschiede zwischen Kindern mit SPF in Förder- und Grundschulen können dadurch nicht nur auf institutionelle und kompositionsbezogene Effekte zurückgeführt werden, sondern auch darauf, dass sich Schülerinnen und Schüler in verschiedenen Schularten systematisch in ihren individuellen Lernausgangsbedingungen unterscheiden. Um diese differenzielle Eingangsselektivität der Schularten zu berücksichtigen, wurde für demografische, schulleistungsrelevante und sozio-kulturelle Merkmale der Schülerinnen und Schüler kontrolliert. Die Auswahl der Kontrollvariablen auf Individualebene orientiert sich an theoretischen Annahmen zum Zuweisungsprozess der Schülerinnen und Schüler mit SPF zu Förder- und Grundschulen als auch an amtlichen Schulstatistiken und empirischen Studien zu sozio-kulturellen, kognitiven und demografischen Unterschieden beider Schülergruppen (Autorengruppe Bildungsberichterstattung 2010, Klein 2001, Kocaj et al. 2014, Kornmann 2006, Wocken 2005).

Insbesondere kognitive Grundfähigkeiten werden als bedeutsame Kontrollvariable für das Vorwissen von Kindern in querschnittlichen Studien zu Schularteffekten betrachtet (Baumert et al. 2006). Vorangegangene Untersuchungen konnten zeigen, dass diese Kontrollvariable zwar nicht Maße des bereichsspezifischen Vorwissens ersetzen kann, jedoch für die Zuweisung zu verschiedenen Schularten prädiktiv ist (Baumert et al. 2006, Kocaj et al. 2014). So unterscheiden sich Schülerinnen und Schüler verschiedener Schularten in ihren kognitiven Fähigkeiten. Allerdings stellt das Merkmal keinen Ersatz für das bereichsspezifische Vorwissen dar, da auch die kognitiven Fähigkeiten durch die schulartspezifische Lernumgebung beeinflusst werden und zeitlichen Veränderungen unterliegen (Becker et al. 2012). Werden die kognitive Grundfähigkeiten der Schülerinnen und Schüler jedoch nicht berücksichtigt, kann dies zu einer Überschätzung der Kompositionseffekte führen. Das liegt daran, dass vorherige Leistungen und individuelle kognitive Fähigkeiten (als Indikator für das Leistungspotenzial) deutlich stärker mit schulischen Kompetenzen zusammenhängen als leistungsbezogene Merkmale auf Klassen- bzw. Schulebene (Köller & Baumert 2001, Köller et al. 2013, Nikolova 2011). In Modellen, in denen das individuelle Vorwissen oder die kognitiven Fähigkeiten nicht berücksichtigt werden, könnten Kompositionsmerkmale Leistungsunterschiede zwischen Schülerinnen und Schülern erklären, die eher auf unterschiedliche individuelle Lernvoraussetzungen zurückführbar sind. Um Aussagen über Zusammenhänge zwischen der Klassenzusammensetzung und individuellen schulischen Kompetenzen treffen zu können, muss deshalb zunächst sichergestellt werden, dass diese nicht primär auf Unterschieden in individuellen Lernausgangsbedingungen basieren (vgl. Opdenakker & van Damme 2001).

4.3.5.1 Stichprobe

Die vorliegende Untersuchung basiert auf Daten des querschnittlich angelegten IQB-Ländervergleichs 2011 für die Primarstufe (Stanat et al. 2012). Die Analysen beziehen sich auf eine Teilstichprobe von Schulklassen, in denen mindestens ein Kind mit einem SPF in den Förderschwerpunkten *Lernen*, *Sprache*, oder *emotionale und soziale Entwicklung* unterrichtet wurde. Schülerinnen und Schüler mit diesen Förderschwerpunkten wurden als Teil der Zielpopulation definiert. Diese Teilstichprobe umfasst insgesamt 8 816 Schülerinnen und Schüler aus 451 Schulklassen (durchschnittliches Alter 10.51 Jahre, $SD = 0.54$, 51.8% Jungen), was einem Anteil von 32.6 Prozent der realisierten Gesamtstichprobe des IQB-Ländervergleichs 2011 entspricht ($N = 27\,081$). 1 195 dieser Schülerinnen und Schüler hatten

einen diagnostizierten SPF in den Bereichen *Lernen*, *Sprache*, oder *emotionale und soziale Entwicklung*. Innerhalb dieser Schülergruppe wurden 762 Kinder in 402 Grundschulen und 433 Kinder in 49 Förderschulen unterrichtet (Tab. 12). Das durchschnittliche Alter der Kinder mit SPF, bei denen es sich zu 66.3 Prozent um Jungen handelte, betrug 10.9 Jahre ($SD = 0.64$). Im Durchschnitt wurden in Förderschulklassen rund 10 Schülerinnen und Schüler und in Grundschulen rund 2 Kinder mit SPF pro Schulklasse getestet (Tab. 12 & Tab. 13). Eine Förderschulklasse wurde von den Analysen ausgeschlossen, da nur ein Schüler getestet werden konnte.

Tabelle 12. *Stichprobengröße nach Beschulungsart*

Schulart	$N_{\text{Kinder mit SPF}}^1$	N_{Klassen}	$M_{\text{Kinder mit SPF}}$	$SD_{\text{Kinder mit SPF}}$	$Min_{\text{Kinder mit SPF}}$	$Max_{\text{Kinder mit SPF}}$	$N_{\text{Kinder insgesamt}}$
Förderschule	433	49	10.24	4.07	2	27	433
Grundschule	762	402	1.90	1.36	1	9	8383

¹sonderpädagogischer Förderbedarf

Tabelle 13. *Klassengröße nach Beschulungsart*

Schulart	N_{Klassen}	$M_{\text{Klassengröße}}$	$SD_{\text{Klassengröße}}$	$Min_{\text{Klassengröße}}$	$Max_{\text{Klassengröße}}$
Förderschule	49	10.24	4.07	2	27
Grundschule	402	20.85	4.15	6	33

In die Analysen gingen nur Kinder mit SPF ein, die an den Kompetenztestungen teilgenommen hatten. Wie in anderen Schulleistungstudien üblich (vgl. OECD 2012, Olson et al. 2008), wurden Schülerinnen und Schüler von der Teilnahme an den Kompetenztests ausgeschlossen, wenn sie eine dauerhafte körperliche Beeinträchtigung aufwiesen, die eine Bearbeitung des Tests unmöglich machte. Auch Schülerinnen und Schüler, die emotional oder geistig nicht in der Lage waren, den allgemeinen Anweisungen der Testung zu folgen, oder die durch die Testsituation in unzumutbarem Maße emotional belastet worden wären, nahmen nicht an der Erhebung teil. Darüber hinaus wurden Schülerinnen und Schüler nicht deutscher Muttersprache ausgeschlossen, die Deutsch weder lesen noch schreiben konnten und weniger als ein Jahr in deutscher Sprache unterrichtet wurden (Richter et al. 2012). Insgesamt wurden 10.4 Prozent der Kinder mit einem SPF in den drei genannten Förderschwerpunkten von den

Analysen ausgeschlossen, wobei von diesen die Mehrheit aufgrund von Abwesenheit bzw. fehlender Elterngenehmigung (6.9%) nicht an der Kompetenztestung teilnahm. Lediglich 3.5 Prozent der Kinder mit SPF wurden aufgrund einer emotionalen bzw. geistigen Beeinträchtigung oder aufgrund fehlender Sprachkenntnisse ausgeschlossen. Ein Ausschluss aufgrund einer dauerhaften körperlichen Beeinträchtigung kam in dieser Schülergruppe nicht vor.

4.3.5.2 Erhebungsinstrumente

Schulische Kompetenzen. Die schulischen Kompetenzen der Schülerinnen und Schüler in Mathematik und im Leseverständnis wurden mit standardisierten Leistungstests erfasst. In beiden Kompetenzbereichen wurden Items mit verschiedenen Aufgabenformaten wie z. B. Multiple-Choice, Richtig-Falsch-Aufgaben, Zuordnungen, Lückenergänzungen und offene Aufgaben bzw. Berechnungen eingesetzt (siehe Stanat et al. 2012). Die Testbedingungen unterschieden sich zwischen Förder- und Grundschulen in der Bearbeitungszeit (80 Minuten pro Kompetenzbereich in Grundschulen, 40 Minuten in Förderschulen), in der Anzahl und im Schwierigkeitsgrad der Aufgaben (weniger und leichtere Aufgaben in Förderschulen). Testitems in Mathematik und im Leseverständnis, die sowohl in Grundschulen als auch in Förderschulen eingesetzt wurden, wiesen in beiden Schülergruppen vergleichbare psychometrische Eigenschaften auf (Kocaj et al. in Vorbereitung). Zur Überprüfung der Vergleichbarkeit der Testwerte von Kindern mit SPF in Förder- und Grundschulen wurden u.a. die Passung des Raschmodells für beide Schülergruppen geprüft, Analysen zum uniformen *Differential Item Functioning (DIF)* durchgeführt und geprüft, ob die Testergebnisse untereinander und mit externen Kriterien in beiden Schülergruppen in vergleichbarer Weise korrelierten. Die resultierenden Testwerte ließen sich auf einer gemeinsamen Skala abbilden. Auch die Befunde der DIF-Analysen und der korrelativen Analysen deuten darauf hin, dass die Kompetenztests in Förder- und Grundschulen ähnlich gut messen und die Testergebnisse der Schülerinnen und Schüler mit SPF für beide Schularten valide verglichen werden können.

Kognitive Grundfähigkeiten. Bei den kognitiven Grundfähigkeiten der Schülerinnen und Schüler handelt es sich um eine zentrale Kontrollvariable für die differenzielle Eingangsselektivität der Förder- und Grundschulen. Zur Erfassung der kognitiven Grundfähigkeiten kamen zwei Subskalen des Kognitiven Fähigkeitstests (KFT 4-12+R, Heller & Perleth 2000) zum Einsatz. Dazu gehörte die Subskala *Wortschatz* (25 Items, $\alpha = .77$) als Indikator für verbale Fähigkeiten und die Subskala *Figurenanalogien* (25 Items, $\alpha = .93$) zur Erfassung nonverba-

ler kognitiver Grundfähigkeiten. Die Anzahl der richtigen Antworten wurde für jede Subskala z-standardisiert. Höhere Ausprägungen auf den Variablen kennzeichnen höhere kognitive Grundfähigkeiten. Als Indikator für das mittlere Fähigkeitsniveau auf Klassenebene wurden anschließend nur die nonverbalen kognitiven Grundfähigkeiten genutzt. Grund für dieses Vorgehen war, dass der Untertest *Figurenanalogien* eine höhere interne Konsistenz aufwies und dadurch einem Aggregationsbias durch unreliable Maße auf Individualebene entgegengewirkt wird, der zu einer Überschätzung von Kompositionseffekten führen kann (vgl. Baumert et al. 2006). Zusätzlich zu den kognitiven Grundfähigkeiten wurden demografische und sozio-kulturelle Merkmale der Schülerinnen und Schüler erhoben, um für bereits bestehende leistungsrelevante Unterschiede zwischen Förder- und Grundschülerinnen und -schülern zu kontrollieren.

Demografische Merkmale. Das Geschlecht und das Alter der Schülerinnen und Schüler wurde mit der Schülerteilnahmeliste erfasst, die von der zuständigen Schulkoordination ausgefüllt wurde. Die Schülerinnen und Schüler machten in den Fragebögen weiterhin Angaben zu ihrem Schuleintrittsalter (in Jahren) und ihre Eltern gaben Auskunft zur Dauer des Besuchs einer Kindertagesstätte bzw. Vorschule ihrer Kinder (*1 = kein Besuch einer Kindertagesstätte oder einer Vorschule, 2 = kürzer als 1 Jahr, 3 = 1-2 Jahre, 4 = 2-3 Jahre, 5 = länger als 3 Jahre*).

Sozio-kultureller Hintergrund. Bei den Merkmalen des sozio-kulturellen Hintergrunds wird allgemein zwischen strukturellen und prozessualen Aspekten familiärer Einflüsse unterschieden (vgl. McElvany et al. 2009). Zu den strukturellen Merkmalen zählen der höchste *International Socio-Economic Index of Occupational Status (HISEI)* des Vaters bzw. der Mutter (Ganzeboom et al. 1992), der höchste Bildungsabschluss der Eltern (*International Standard Classification of Education, HISCED*, OECD 1999) und die berufliche Situation der Eltern. Der höchste Bildungsabschluss der Eltern kennzeichnet das Bildungsniveau in der Familie und basiert auf Informationen der Eltern zum höchsten Schul- und Berufsabschluss. Höhere Werte auf dieser sechsstufigen Ratingskala entsprechen einem höherwertigen Abschluss. Die Erfassung der beruflichen Situation basiert auf Angaben der Eltern, ob sie zurzeit bezahlt in Vollzeit arbeiten, eine Teilzeittätigkeit ausüben oder keiner bezahlten Arbeit nachgehen. Zur Feststellung eines Zuwanderungshintergrunds in der Familie wurde auf Angaben aus dem Elternfragebogen zurückgegriffen. Ein Zuwanderungshintergrund lag vor, wenn mindestens ein Elternteil nicht in Deutschland geboren wurde. Weiterhin gingen Informationen zur in der Familie gesprochenen Sprache als Kontrollvariable in die Analysen ein. Dabei wurde

unterschieden, ob immer bzw. fast immer Deutsch oder manchmal bzw. nie Deutsch zu Hause gesprochen wird.

Ein wichtiger prozessualer Aspekt des sozio-kulturellen Hintergrunds ist der häusliche Buchbestand. Die Schülerinnen und Schüler wurden anhand eines Items gefragt, wie viele Bücher bei Ihnen zu Hause vorhanden sind (Bücher zu Hause). Die Antwortskala bestand aus 5 Stufen [*1 = keine oder nur sehr wenige (0-10 Bücher), 2 = genug, um ein Regalbrett zu füllen (11-25 Bücher), 3 = genug, um ein Regal zu füllen (26-100 Bücher), 4 = genug, um zwei Regale zu füllen (101-200 Bücher), 5 = genug, um drei oder mehr Regale zu füllen (über 200 Bücher)*]. Daneben wurde die durchschnittliche Lesezeit als weiteres prozessuales Merkmal des sozio-kulturellen Hintergrunds erfasst. Die Schülerangaben zur Lesezeit bezogen sich darauf, wie viel Zeit diese außerhalb der Schule an einem normalen Schultag mit Lesen verbringen. Hier wurde differenziert zwischen *(1) weniger als 30 Minuten, (2) 30 Minuten bis zu einer Stunde, (3) ein bis zwei Stunden und (4) zwei Stunden oder mehr pro Tag*. Außerdem schätzten die Eltern die Bedeutsamkeit der Vermittlung vorgegebener Erziehungsziele in der Familie ein. Die Skala (Baumert et al. 2008) beschreibt, für wie wichtig Eltern die Vermittlung von Erziehungszielen wie „solide Schulkenntnisse“ oder „persönliche Selbstständigkeit“ erachten. Hohe Werte auf der aggregierten Skala (18 Items, $\alpha = .90$) weisen darauf hin, dass diese Erziehungsziele für Eltern eine hohe Bedeutung haben. Als weiteres Merkmal des sozio-kulturellen Hintergrunds wurden die Bildungsaspirationen der Eltern (für das Kind gewünschter bzw. erwarteter Schulabschluss) betrachtet. Die Erfassung der Bildungsaspirationen der Eltern basiert in der vorliegenden Untersuchung auf zwei Items. Zum einen wurde anhand einer sechsstufigen Ratingskala (*1 = überhaupt nicht, 6 = voll und ganz*) erfragt, wie sehr sich die Eltern einen Hauptschulabschluss für ihr Kind wünschen. Außerdem sollten die Eltern auf einer vierstufigen Ratingskala (*1 = überhaupt nicht wahrscheinlich, 4 = sehr wahrscheinlich*) angeben, für wie wahrscheinlich ein Real-schulabschluss des Kindes erachtet wird.

Teilleistungsstörungen. Darüber hinaus gingen Informationen der Schulkoordination zu diagnostizierten Teilleistungsstörungen der Kinder mit SPF in die Analysen ein. Diese Angaben beinhalten die Kategorien *(1) Aufmerksamkeitsdefizit-Hyperaktivitätsstörung (ADHS), (2) Lernstörungen (isolierte Rechtschreibstörung, Lese-Rechtschreibstörung, Rechenstörung), (3) Mehrfachdiagnose (Kombination aus ADHS und einer Lernstörung) sowie (4) nicht definierte Teilleistungsstörungen.*

Darüber hinaus wurde das Bundesland als Kontrollvariable in den Analysen berücksichtigt, da zwischen den Ländern deutliche Unterschiede im Anteil der Schülerinnen und Schüler mit SPF und deren Verteilung auf allgemeine Schulen und Förderschulen bestehen (Autorengruppe Bildungsberichterstattung 2014).

4.3.5.3 Datenanalyse

Zur Bestimmung institutioneller und kompositionsbezogener Effekte wurden eine Reihe von Mehrebenenanalysen (*Random-Intercept-Modelle*, Raudenbush & Bryk 2002) in Mplus 7.0 (Muthén & Muthén 1998-2012) durchgeführt. Dabei wurden Modelle spezifiziert, in die sukzessive Prädiktoren auf Individual- und Klassenebene aufgenommen wurden. Fehlende Werte in den betrachteten Merkmalen wurden mehrfach imputiert und durch jeweils 15 plausible Ausprägungen ersetzt (van Buuren & Groothuis-Oudshoorn 2011).

Auf Klassenebene basieren die Kompositionsmerkmale auf den klassenweise aggregierten Werten des HISEI und der nonverbalen kognitiven Grundfähigkeiten. Diese wurden nach der Aggregation z-standardisiert. In Grundschulen wurden bei der Bildung dieser beiden Kompositionsmerkmale alle Kinder einer Klasse einbezogen, in die Mehrebenenanalysen anschließend jedoch nur Kinder mit einem SPF aufgenommen. Durch die geringe Anzahl der Kinder mit SPF pro Klasse in Grundschulen ist bei der Schätzung der Regressionsgewichte auf Individualebene mit einer Erhöhung der Standardfehler zu rechnen (Hox 2002). Außerdem kann auch die hohe Anzahl an Prädiktoren auf Individualebene, die der Kontrolle der differenziellen Eingangsselektivität der Schularten dienen, zu einer Erhöhung der Standardfehler führen (Schafer & Kang 2008). Die Schätzungen der Kompositionseffekte und institutionellen Effekte auf Klassenebene sollten durch die geringe Clustergröße in Grundschulen bis auf eine Verringerung der Teststärke jedoch nicht verzerrt sein, da für diese Analysen die Anzahl der Schulklassen der ausschlaggebende Faktor ist (Bell et al. 2010, Hox 2002, Snijders & Bosker 2012).

4.3.6 Ergebnisse

4.3.6.1 Demografische und sozio-kulturelle Unterschiede zwischen Kindern mit SPF in Förder- und Grundschulen

In der ersten Fragestellung sollte geprüft werden, ob Schülerinnen und Schüler mit SPF in Grundschulen über einen günstigeren sozio-kulturellen Hintergrund verfügen als Kinder in Förderschulen. Außerdem wurde geprüft, ob auch in demografischen Merkmalen Unterschiede zwischen Schülerinnen und Schülern mit SPF in Förder- und Grundschulen bestehen (vgl. Kocaj et al. 2014). Die Ergebnisse des Gruppenvergleichs für metrisch skalierte Merkmale sind in Tabelle 14 dargestellt, die Befunde zu den nominalen Variablen in Tabelle 15.

Tabelle 14. *Demografische und sozio-kulturelle Unterschiede zwischen Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen: metrisch skalierte Merkmale*

	Förderschule		Grundschule		<i>t</i> (1069)	<i>p</i>	Cohens <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Alter (in Jahren)	10.97	0.61	10.86	0.64	2.96	.003	−0.18
Schuleintrittsalter (in Jahren)	6.48	0.75	6.21	0.67	6.24	<.001	−0.39
Höchster Bildungsabschluss in der Familie (HISCED ¹)	3.10	1.13	3.57	1.20	−6.39	<.001	0.40
Sozialer Status (HISEI ²)	36.79	13.27	41.88	14.57	−5.76	<.001	0.36
Voraussichtlicher Schulabschluss: mindestens RSA ³	2.75	0.90	3.20	0.78	−8.57	<.001	0.54
Gewünschter Schulabschluss: HSA ⁴	4.22	1.90	3.26	2.01	7.78	<.001	−0.49
Bücher zu Hause	2.51	1.24	2.82	1.23	−4.00	<.001	0.25
Lesezeit	1.79	0.97	2.06	1.09	−4.09	<.001	0.26

N = 1071. Das nominale α -Niveau beträgt nach Anwendung der Bonferroni-Korrektur $p = .004$ (13 Vergleiche bei einer jeweiligen Irrtumswahrscheinlichkeit von 5% entspricht $.05/13 = .00385$). ¹ Highest International Standard Classification of Education, ² Highest International Socio-Economic Index of Occupational Status, ³ Realschulabschluss, ⁴ Hauptschulabschluss

Demografische Merkmale. Zunächst wurde geprüft, ob sich Kinder mit SPF in Förderschulen und Grundschulen in den untersuchten demografischen Merkmalen voneinander unterscheiden. Kein bedeutsamer Unterschied zeigte sich in der Geschlechterverteilung [$\chi^2(1) = 0.00$, $p = 1.00$]. Sowohl in Förderschulen als auch in Grundschulen waren Jungen in der Schüler-

gruppe mit SPF überrepräsentiert. Schülerinnen und Schülern mit SPF in Förderschulen waren etwas älter als Kinder mit SPF in Grundschulen [$t(1069) = 2.96, p = .003, d = -0.18$]. Des Weiteren wurden Kinder mit SPF in Grundschulen durchschnittlich früher eingeschult als Kinder, die eine Förderschule besuchten [$t(1069) = 6.24, p < .001, d = -0.39$].

Tabelle 15. *Demografische und sozio-kulturelle Unterschiede zwischen Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen: nominal skalierte Merkmale*

	Förderschule		Grundschule		χ^2	df	p	OR ¹
	absolute Häufigkeit	relative Häufigkeit	absolute Häufigkeit	relative Häufigkeit				
Geschlecht								
männlich	278	67.3%	443	67.3%	0.00	1	1.00	1.00
weiblich	135	32.7%	215	32.7%				
Zuwanderungshintergrund der Eltern								
beide Eltern in Deutschland geboren	320	77.5%	497	75.5%	0.53	1	.51	1.11
ein oder beide Elternteil(e) im Ausland geboren	93	22.5%	161	24.5%				
Sprache zu Hause								
Kind spricht manchmal bzw. nie Deutsch	32	7.7%	50	7.6%	0.01	1	1.00	1.02
Kind spricht immer oder fast immer Deutsch	381	92.3%	608	92.4%				
Berufliche Situation der Eltern: Vater								
arbeitet bezahlt in Vollzeit	226	54.7%	529	80.4%	84.62	4	<.001	0.29
arbeitet bezahlt in Teilzeit	28	6.8%	16	2.4%				2.92
hat keine bezahlte Arbeit	63	15.3%	33	5.0%				3.41
anderes	56	13.6%	51	7.8%				
nichts trifft zu	40	9.7%	29	4.4%				
Berufliche Situation der Eltern: Mutter								
arbeitet bezahlt in Vollzeit	98	23.7%	166	25.2%	15.80	4	.003	0.92
arbeitet bezahlt in Teilzeit	121	29.3%	246	37.4%				0.69
hat keine bezahlte Arbeit	81	19.6%	123	18.7%				1.06
anderes	58	14.0%	74	11.2%				
nichts trifft zu	55	13.3%	49	7.4%				

$N = 1071$. Das nominale α -Niveau beträgt nach Anwendung der Bonferroni-Korrektur $p = .004$ (13 Vergleiche bei einer jeweiligen Irrtumswahrscheinlichkeit von 5% entspricht $.05/13 = .00385$)

Sozio-kulturelle Merkmale. Für Schülerinnen und Schüler mit SPF zeigten sich deutliche Unterschiede im sozio-kulturellen Hintergrund zwischen Förder- und Grundschulen. Kinder mit SPF in Grundschulen scheinen im Durchschnitt in günstigeren sozio-kulturellen Verhältnissen aufzuwachsen als Kinder mit SPF in Förderschulen (Tab. 14 & Tab. 15). Dies zeigte

sich sowohl für strukturelle Merkmale (berufliche Situation der Eltern, sozialer Status, höchster Bildungsabschluss der Eltern) als auch für prozessuale Aspekte (Buchbestand in der Familie, durchschnittliche Lesezeit der Kinder außerhalb der Schule) des sozio-kulturellen Hintergrundes und für die Bildungsaspirationen der Eltern. Bei den Ausprägungen der Unterschiede handelt es sich nach Cohen (1992) um kleine bis mittlere Effekte. Im Gegensatz dazu konnten beim Vergleich der beiden Schularten keine Unterschiede für den Zuwanderungshintergrund der Eltern und der zu Hause gesprochenen Sprache festgestellt werden.

4.3.6.2 Zusammenhänge der Klassenzusammensetzung und der Beschulungsart mit den schulischen Kompetenzen von Kindern mit SPF

Ziel der Mehrebenenanalysen war es zu prüfen, ob institutionelle Merkmale und Merkmale der Klassenkomposition über individuelle Schülermerkmale hinaus mit den schulischen Kompetenzen von Schülerinnen und Schülern mit SPF zusammenhängen. Als Kompositionsmerkmale wurden die soziale und fähigkeitsbezogene Klassenzusammensetzung und ihre Beziehung mit den Kompetenzständen in Mathematik und im Leseverständnis betrachtet.

Deskriptive Befunde. In einem ersten Schritt wurde der Zusammenhang zwischen sozialer Herkunft und nonverbalen kognitiven Grundfähigkeiten auf Klassenebene in Förder- und Grundschulen betrachtet. Diese Beziehung wird als Indikator für die soziale Trennung der Schülerschaft durch die Zuweisung zu verschiedenen Schularten verwendet. In Abbildung 8 wird ersichtlich, dass ein mittelstarker positiver Zusammenhang zwischen sozialer Herkunft und Fähigkeitsniveau der Schulklassen besteht. 32 Prozent der Varianz in den mittleren nonverbalen kognitiven Fähigkeiten ließen sich durch Unterschiede in der sozialen Komposition der Schülerschaft beschreiben.

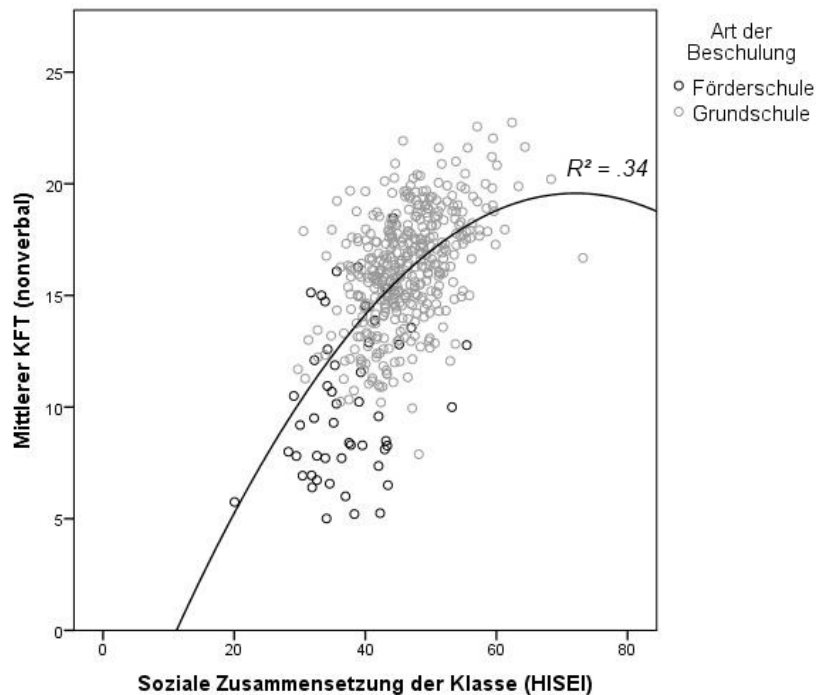


Abbildung 8. Zusammenhang zwischen sozialer Zusammensetzung (HISEI) und mittlerem Niveau der nonverbalen kognitiven Grundfähigkeiten (KFT) auf Klassenebene (linearer und quadratischer Trend).

Außerdem wurden korrelative Zusammenhänge aller berücksichtigten Merkmale auf Klassenebene bestimmt (Tab. 16). Es zeigten sich positive Zusammenhänge mittlerer Stärke zwischen der Beschulungsart und der sozialen ($r = .41, p < .001$) und fähigkeitsbezogenen Klassenzusammensetzung ($r = .58, p < .001$). Förderschulklassen wiesen im Durchschnitt ein niedrigeres kognitives Fähigkeitsniveau und eine ungünstigere soziale Zusammensetzung auf als Grundschulklassen (Tab. 17). Allerdings hatten einige Förderschulklassen vergleichbare mittlere kognitive Fähigkeiten bzw. eine vergleichbare soziale Zusammensetzung wie Grundschulklassen. Auch eine grafische Inspektion der Verteilungen der Kompositionsmerkmale weist auf einen hinreichenden Überlappungsbereich der Kompositionsmerkmale zwischen Förder- und Grundschulen hin (Anhang A1 & Anhang A2).

Tabelle 16. *Korrelationen zwischen Art der Beschulung und Merkmalen der Klassenzusammensetzung*

	1	2	3	4
1 Grundschule				
2 Mittlere Leseleistung (PVs ¹)	.40***			
3 Mittlere Mathematikleistung (PVs)	.44***	.97***		
4 Mittlere kognitive Grundfähigkeiten (nonverbal)	.58***	.47***	.56***	
5 Mittlere soziale Zusammensetzung (HISEI ²)	.41***	.36***	.40***	.58***

* $p < .05$, ** $p < .01$, *** $p < .001$. ¹ Plausible Values, ² Highest International Socio-Economic Index of Occupational Status

Tabelle 17. *Deskriptive Kennwerte der mittleren Klassenzusammensetzung in Förder- und Grundschulen*

	N_{Klassen}	M	SD	Min	Max
Mittlere kognitive Grundfähigkeiten (nonverbal)					
Förderschule	49	10.06	3.42	5.01	18.46
Grundschule	402	16.22	2.54	7.89	22.74
Mittlere soziale Zusammensetzung (HISEI ¹)					
Förderschule	49	36.89	5.98	20.06	55.46
Grundschule	402	46.29	6.11	29.73	73.20

¹ Highest International Socio-Economic Index of Occupational Status

Mehrebenenanalysen zum Zusammenhang zwischen Klassenkomposition, Beschulungsart und schulischen Kompetenzen. Um die Zusammenhänge der Beschulungsart und Komposition separat, aber auch im Zusammenspiel analysieren zu können, erfolgten die Mehrebenenanalysen mehrstufig. In einem ersten Schritt (Modell 1, Tab. 18 zur Vorhersage der Mathematikleistung und Tab. 19 zur Vorhersage der Leseverständnisleistung) wurde ein Modell ohne Prädiktoren geprüft, um der Frage nachzugehen, wie stark die Testleistungen von Kindern mit SPF zwischen und innerhalb von Schulklassen variieren. Die resultierende Intraklassenkorrelation (*ICC*) von .59 für Mathematik indiziert, dass 59 Prozent der beobachteten Leistungsunterschiede zwischen Kindern mit SPF durch Unterschiede zwischen Schulklassen beschrieben werden können. Für die Leistungsunterschiede im Leseverständnis war die *ICC* mit .53 nur geringfügig niedriger. Um die Eingangsselektivität in Förder- und Grundschulen berücksichtigen zu können, wurden in Modell 2 individuelle Merkmale der Kinder mit SPF zur Vorhersage ihrer Testleistungen in Mathematik und im Leseverständnis aufgenommen (Tab. 18 bzw.

19). Dabei waren erwartungsgemäß die kognitiven Grundfähigkeiten für beide Kompetenzbereiche jeweils die stärksten Prädiktoren. Modell 3 enthält zusätzlich die Beschulungsart als unabhängige Variable (Tab. 18 bzw. 19), um zu prüfen, ob der Besuch einer Förder- oder Grundschule nach Kontrolle individueller Merkmale mit den schulischen Kompetenzen der Kinder mit SPF zusammenhängt. Die Ergebnisse zeigen, dass Kinder mit SPF in Grundschulen nach Kontrolle individueller Merkmale signifikant höhere schulische Kompetenzen aufwiesen als Kinder in Förderschulen.

Um den Effekten der Klassenkomposition nachzugehen, wurde statt der Beschulungsart das mittlere kognitive Fähigkeitsniveau (Modell 4 in Tab. 18 & 19) bzw. die soziale Zusammensetzung (Modell 5 in Tab. 18 & 19) als Prädiktoren auf Klassenebene aufgenommen. Dabei ließen sich positive Zusammenhänge der Klassenzusammensetzung mit den schulischen Kompetenzen von Schülerinnen und Schülern mit SPF feststellen. In Klassen mit hohen mittleren kognitiven Fähigkeiten und in Klassen mit einer günstigen sozialen Zusammensetzung erzielten Kinder mit SPF höhere Leistungen in Mathematik und im Leseverständnis.

In den nächsten Modellen wurden das mittlere Fähigkeitsniveau und die soziale Zusammensetzung der Klasse gemeinsam mit der Schulart berücksichtigt, um die spezifischen Zusammenhänge dieser Merkmale mit den Testleistungen zu bestimmen. Bei simultaner Berücksichtigung der mittleren kognitiven Fähigkeiten auf Klassenebene und der Schulart gingen zwar beide Effekte zurück, blieben aber von substanzieller Bedeutung (Modell 6, Tab. 18 & 19). In Modell 7 wurde die soziale Zusammensetzung der Schülerschaft neben der Beschulungsart aufgenommen. Der Zusammenhang zwischen der sozialen Zusammensetzung und der Testleistung ging im Vergleich zum Modell 5 zurück, blieb für Mathematik aber bedeutsam (Modell 7, Tab. 18). Im Leseverständnis hingegen hing die soziale Zusammensetzung nicht mehr signifikant mit den Testleistungen zusammen, wenn gleichzeitig auch die Beschulungsart berücksichtigt wurde (Modell 7, Tab. 19). In Modell 8 zeigte sich dann jedoch, dass bei simultaner Berücksichtigung der Beschulungsart und der mittleren kognitiven Fähigkeiten der Klasse die soziale Zusammensetzung der Schülerschaft keinen bedeutsamen Beitrag zur Vorhersage von Leistungsunterschieden zwischen Kindern mit SPF leistet (Tab. 18 & 19). Bei der gleichzeitigen Aufnahme aller betrachteten Prädiktoren auf Klassenebene ergab sich außerdem ein differenzielles Befundmuster für Mathematik und für das Leseverständnis (Modell 8 in Tab. 18 & 19). Im Leseverständnis verringerten sich die Koeffizienten für die Beschulungsart und das kognitive Fähigkeitsniveau der Klasse zwar, wenn institutionelle und kompositionsbezogene Merkmale gleichzeitig berücksichtigt wurden, sie blieben aber beide

von signifikanter Bedeutung für die Vorhersage von Leistungsunterschieden (Modell 8, Tab. 19). In Mathematik führte die gleichzeitige Berücksichtigung institutioneller und kompositioneller Merkmale dazu, dass nur noch das mittlere kognitive Fähigkeitsniveau signifikant zur Vorhersage von Leistungsunterschieden beitrug (Modell 8, Tab. 18). Der Zusammenhang der Beschulungsart mit den Testleistungen in Mathematik war statistisch nicht mehr bedeutsam, verfehlte die Signifikanzschwelle jedoch nur knapp ($b = 23.25$, $p = .055$). Das könnte auf die höheren Standardfehler im Kompetenzbereich Mathematik aufgrund der geringeren Anzahl an Schulklassen zurückgeführt werden. Davon abgesehen war die der Einfluss der Beschulungsart unter Berücksichtigung des Fähigkeitsniveaus und der sozialen Zusammensetzung der Klasse in beiden Kompetenzbereichen von vergleichbarer Stärke.

Zusammenfassend konnte die Hypothese bestätigt werden, dass Merkmale der Klassenkomposition nach Kontrolle individueller Lernvoraussetzungen bedeutsam zur Vorhersage von Leistungsunterschieden zwischen Kindern mit SPF in Förder- und Grundschulen beitragen. Institutionelle Effekte der Beschulungsart gingen nach Kontrolle kompositionsbezogener Merkmale deutlich zurück. Dabei war das mittlere Fähigkeitsniveau das bedeutsamste Kompositionsmerkmal. In Schulklassen, in denen das Fähigkeitsniveau eine Standardabweichung über dem Durchschnitt lag, erhöhte sich die vorausgesagte Testleistung der Schülerinnen und Schüler mit SPF in Mathematik um ca. 19 Punkte und im Leseverständnis um ca. 10 Punkte. Bezogen auf die Metrik der Bildungsstandards ($M = 500$, $SD = 100$) entsprechen diese Unterschiede 0.19 Standardabweichungen in Mathematik und 0.10 Standardabweichungen im Leseverständnis.

Tabelle 18. Mehrebenenmodell zur Vorhersage der Mathematikleistung durch individuelle Schülermerkmale, Beschulungsart und Klassenkomposition ($N = 1052$ Schülerinnen und Schüler aus 418 Klassen)

	Testleistung Mathematik															
	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6	Modell 7	Modell 8								
	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	560.75***	79.03	483.81***	75.38	588.31***	74.35	546.18***	77.11	552.95***	75.74	492.65***	74.92	551.05***	75.62		
Individualebene (L1)																
Geschlecht (1 = weiblich)	-29.41***	4.77	-30.70***	4.72	-31.84***	4.81	-29.91***	4.70	-31.88***	4.78	-30.75***	4.70	-31.84***	4.77		
Alter	-11.70**	4.43	-12.01**	4.35	-12.12**	4.29	-11.25*	4.38	-12.16**	4.27	-11.79**	4.32	-12.08**	4.27		
Teilleistungsstörung	0.83	2.70	-0.57	2.59	-0.57	2.53	-0.44	2.66	-0.88	2.53	-1.01	2.60	-0.96	2.53		
Bildung der Eltern (HISCED) ¹	0.52	2.65	0.32	2.59	0.06	2.59	0.32	2.63	0.06	2.58	0.23	2.59	0.05	2.59		
Sozialer Status (HISEI) ²	1.30	3.58	0.88	3.55	0.42	3.59	-1.07	3.60	0.41	3.58	-0.36	3.59	0.15	3.65		
Vermittlung von Erziehungszielen	-1.96	5.30	-1.58	5.27	-1.94	5.23	-1.61	5.25	-1.79	5.24	-1.50	5.25	-1.75	5.24		
Voraussichtlicher Schulabschluss: mind. RSA ³	9.19*	4.07	8.20*	4.08	8.10*	4.03	8.53*	4.00	7.87	4.05	7.97	4.04	7.84	4.04		
Gewünschter Schulabschluss: HSA ⁴	-1.66	1.37	-1.10	1.36	-1.08	1.38	-1.28	1.36	-0.95	1.37	-0.98	1.36	-0.93	1.37		
Zuwanderungshintergrund ⁵	6.26	7.58	7.23	7.38	7.02	7.40	6.10	7.55	7.27	7.35	7.00	7.39	7.22	7.37		
Sprache zu Hause ⁶	5.75	12.33	6.68	12.23	4.83	12.09	6.07	12.20	5.41	12.08	6.67	12.16	5.47	12.08		
Besuchsdauer Kindergarten	2.87	2.93	2.64	2.96	2.17	2.93	2.56	2.92	2.22	2.94	2.51	2.94	2.22	2.94		
Schuleintrittsalter	-3.26	3.89	-1.72	3.76	-1.62	3.76	-2.58	3.79	-1.28	3.74	-1.61	3.72	-1.28	3.73		
Bücher zu Hause	3.18	2.55	3.32	2.54	3.19	2.52	2.91	2.55	3.25	2.53	3.14	2.54	3.22	2.53		
Kognitive Grundfähigkeiten: nonverbal	30.90***	2.95	30.26***	2.89	26.58***	2.93	30.33***	2.93	27.20***	2.91	30.05***	2.88	27.30***	2.93		
Kognitive Grundfähigkeiten: verbal	40.45***	4.78	38.14***	4.41	37.94***	4.39	39.33***	4.57	37.41***	4.31	37.91***	4.38	37.39***	4.31		
Klassenebene (L2)																
Grundschule (Referenz)																
Förderschule			63.29***	10.51					24.13*	12.17	50.75***	11.39	23.25	12.12		
Mittlerer KFT ⁷ (nonverbal)					25.07***	3.02			20.04***	3.93			19.17***	4.38		
Mittlerer HISEI							18.02***	3.62			9.70**	3.73	2.03	3.94		
Varianzkomponenten																
Individualebene (L1)	5492.65	3947.15	3920.20		4026.72		3928.52		3987.16		3903.47		3981.65			
Klassenebene (L2)	7905.90	2630.35	2149.02		1670.50		2325.60		1699.18		2105.95		1705.33			
Erklärte Varianz auf Individualebene (L1)	28.14%		28.63%		26.69%		28.48%		27.41%		28.93%		27.51%			
Erklärte Varianz auf Klassenebene (L2)	66.73%		72.82%		78.87%		70.58%		78.51%		73.36%		78.43%			

* $p < .05$, ** $p < .01$, *** $p < .001$, ICC = .59. ¹Highest International Standard Classification of Education, ²Highest International Socio-Economic Index of Occupational Status, ³Realschulabschluss, ⁴Hauptschulabschluss, ⁵1 = mindestens ein Elternteil im Ausland geboren, ⁶1 = immer oder fast immer Deutsch, ⁷Kognitive Grundfähigkeiten

4.3.7 Diskussion

In der vorliegenden Studie konnte erstmals anhand einer großen bundesweiten Stichprobe gezeigt werden, dass die Zusammenhänge der Beschulungsart mit den schulischen Kompetenzen von Kindern mit SPF auch auf Merkmale der Klassenzusammensetzung zurückgeführt werden können. Im Rahmen der ersten Fragestellung konnte gezeigt werden, dass Schülerinnen und Schüler mit SPF in Förderschulen hinsichtlich der sozio-kulturellen Zusammensetzung der Schülerschaft ein ungünstigeres Umfeld antreffen als integrativ bzw. inklusiv beschulte Kinder mit SPF. Das spiegelte sich in strukturellen Aspekten (berufliche Situation der Eltern, sozialer Status, höchster Bildungsabschluss der Eltern) als auch prozessualen Merkmalen (häuslicher Buchbestand, Lesezeit) der sozio-kulturellen Herkunft wider. Beide Merkmalsbereiche der sozio-kulturellen Herkunft können sich im Sinne primärer Effekte sozialer Disparitäten auf die schulischen Leistungen der Kinder auswirken (McElvany et al. 2009). Auffallend waren auch die deutlich höheren Bildungsaspirationen der Eltern von Kindern mit SPF in allgemeinen Grundschulen. Diese könnten nach Boudon (1974) als sekundäre Effekte sozialer Disparitäten wirken und die Bildungsentscheidungen der Eltern beeinflussen. Einschränkend muss beachtet werden, dass die Überweisung auf eine Förderschule nicht nur durch die Elternentscheidung, sondern u.a. auch durch den Schweregrad der sonderpädagogischen Förderbedürftigkeit des Kindes beeinflusst wird (zum Prozess der Feststellung eines sonderpädagogischen Förderbedarfs siehe Sälzer et al. 2014).

Die Ergebnisse zur zweiten Fragestellung deuten darauf hin, dass Schülerinnen und Schüler mit SPF in Grundschulen höhere schulische Kompetenzen in Mathematik und im Leseverständnis aufweisen als Kinder in Förderschulen (vgl. auch Kocaj et al. 2014). Die Zusammenhänge der Beschulungsart mit den schulischen Kompetenzen der Kinder mit SPF gingen jedoch deutlich zurück, wenn Merkmale der Klassenkomposition berücksichtigt wurden. Nach Kontrolle der differenziellen Eingangsselektivität der Schularten und der Beschulungsart ließ sich ein positiver Zusammenhang zwischen dem mittleren Fähigkeitsniveau der Schulklassen und den schulischen Kompetenzen von Kindern mit SPF in Mathematik und im Leseverständnis feststellen. Die soziale Klassenzusammensetzung hatte darüber hinaus keine zusätzliche Vorhersagekraft. Diese Befunde sprechen dagegen, dass die Leistungsunterschiede zwischen Schülerinnen und Schülern mit SPF in Förder- und Grundschulen ausschließlich auf institutionelle Effekte (Unterschiede in der Unterrichtsorganisation und in der Unterrichtsgestaltung) zurückführbar sind. Vielmehr trägt auch die fähigkeitsbezogene Klassenzu-

sammensetzung zu Leistungsunterschieden zwischen Schülerinnen und Schülern mit SPF in Förder- und Grundschulen bei. Allerdings ist das mittlere kognitive Fähigkeitsniveau in Grundschulen deutlich höher als in Förderschulen, so dass eine vollständige Trennung von institutionellen und kompositionellen Effekten nicht möglich ist (Dumont et al. 2013, Köller & Baumert 2001).

Bei der Interpretation der vorliegenden Befunde stellt die querschnittliche Anlage der Studie eine zentrale methodische Herausforderung dar. So kann von einem korrelativen Zusammenhang zwischen der Klassenkomposition und individuellen schulischen Kompetenzen nicht auf einen kausalen Wirkmechanismus geschlossen werden. Außerdem konnte die differenzielle Eingangsselektivität der Schularten in der vorliegenden Untersuchung nur begrenzt berücksichtigt werden. Es lagen weder Informationen zum Schweregrad der Beeinträchtigungen noch zum bereichsspezifischen Vorwissen der Schülerinnen und Schüler vor. Zwar standen die kognitiven Grundfähigkeiten als ein Indikator für das Fähigkeitspotenzial der Schülerinnen und Schüler zur Verfügung; dieser Indikator kann jedoch das eigentlich interessierende bereichsspezifische Vorwissen nicht vollständig abbilden. Zudem können kognitive Grundfähigkeiten ihrerseits durch die Schulart und die Klassenkomposition beeinflusst werden (Baumert et al. 2006, Becker et al. 2012). Sollten Grundschulen im Vergleich zu Förderschulen einen stärker positiven Einfluss auf die Entwicklung der kognitiven Grundfähigkeiten haben, dann würde die Aufnahme der kognitiven Grundfähigkeiten im Individualmodell dazu führen, dass dieser Anteil bei der Schätzung des Effekts der Beschulungsart auf schulische Kompetenzen unberücksichtigt bleibt. Diese „Überkontrolle“ auf Individualebene hätte eine Unterschätzung der Kompositionseffekte zur Folge, insofern die Entwicklung der kognitiven Fähigkeiten durch die Zusammensetzung der Schülerschaft beeinflusst wird (*Mediationsbias*, vgl. Baumert et al. 2006).

Allerdings kann die Nichtberücksichtigung bedeutsamer Individualmerkmale in querschnittlichen Studien auch zu einer Überschätzung von Kompositionseffekten führen (*Endogenitätsproblem*, Baumert et al. 2006, Dumont et al. 2013). Deshalb sollten die vorliegenden Ergebnisse in künftigen Studien um Indikatoren des Vorwissens und des Schweregrades der Beeinträchtigung ergänzt und längsschnittlich angelegt werden, um weitere empirische Evidenz zum Einfluss der Beschulungsart und Klassenkomposition auf die Schulleistung von Kindern mit SPF zu gewinnen. Eine weitere Limitation der vorliegenden Untersuchung besteht darin, dass insbesondere in Förderschulen nicht für alle Schülerinnen und Schüler einer Klasse Informationen vorliegen. So ging in die Analysen eine Förderschulklasse ein, in

der nur zwei Kinder getestet wurden (Tab. 13). Falls der Ausfall an Schülerinnen und Schüler pro Klasse nicht zufällig zustande kam, sondern Kinder mit SPF mit besseren kognitiven Fähigkeiten mit höherer Wahrscheinlichkeit am Test teilnahmen, könnte es zu einer Überschätzung der Kompositionseffekte in Förderschulen gekommen sein. Weiterhin waren in der vorliegenden Untersuchung Effekte der Klassenkomposition mit Effekten auf Schulebene konfundiert. Im IQB-Ländervergleich 2011 wurde pro Schule je eine Klasse in die Stichprobe aufgenommen. Dadurch ist es nicht möglich, Merkmale des Lernmilieus auf Klassen- und Schulebene zu differenzieren. Der Fokus der Studie lag jedoch auf Zusammenhänge der Komposition der Schülerschaft mit den schulischen Kompetenzen der Kinder mit SPF. Dabei scheint es bedeutsamer, die Komposition der unmittelbaren Lernumgebung der Schülerinnen und Schüler auf Klassenebene anstatt auf Schulebene zu betrachten (Wilkinson 2002).

In zukünftigen Studien sollte geprüft werden, ob mit Interaktionseffekten zwischen der Beschulungsart und Merkmalen der Klassenzusammensetzung auf die schulischen Kompetenzen zu rechnen ist. So könnte die fähigkeitsbezogene Klassenzusammensetzung in Förder- und Grundschulen unterschiedlich mit den individuellen schulischen Kompetenzen der Schülerinnen und Schüler zusammenhängen. Empirische Untersuchungen in der Sekundarstufe weisen darauf hin, dass sich eine ungünstige fähigkeitsbezogene Komposition vor allem in Hauptschulen negativ auf die schulischen Kompetenzen auswirkt (Baumert et al. 2006). Im Gegensatz dazu fanden Baumert und Kollegen (2006) für Gymnasien keinen Zusammenhang zwischen dem mittleren Fähigkeitsniveau einer Schule und den individuellen schulischen Kompetenzen. Außerdem sollte eine differenzielle Betrachtung nach verschiedenen Förderschwerpunkten vorgenommen werden. In der vorliegenden Untersuchung konnte aufgrund der geringen Anzahl von Schulklassen nicht untersucht werden, ob Merkmale der Klassenkomposition bei Kindern mit den Förderschwerpunkten *Lernen*, *Sprache*, und *emotionale und soziale Entwicklung* unterschiedlich mit den schulischen Kompetenzen zusammenhängen. Ein weiteres Forschungsdesiderat besteht in der direkten Untersuchung vermittelnder Prozesse (z. B. normative Kultur der Schülerschaft, Unterrichtsorganisation, Lehrerwartungen), die den Einfluss von Kompositionsmerkmalen erklären könnten (z.B. Rjosk et al. 2014).

Zusammenfassend liefert die vorliegende Untersuchung erste Hinweise darauf, dass Schülerinnen und Schüler mit SPF in Grundschulen institutionelle und kompositionelle Bedingungen vorfinden könnten, die ihre schulischen Kompetenzen stärker fördern als die Lernumgebung in Förderschulen (vgl. auch Kocaj et al. 2014). Dabei scheint das mittlere Fähigkeitsniveau für die schulischen Kompetenzen der Kinder mit SPF bedeutsamer zu sein als die soziale

Zusammensetzung der Klasse. Die gemeinsame Beschulung von Kindern mit SPF in Grundschulen könnte einen Beitrag dazu leisten, Bildungsbenachteiligungen von Kindern mit SPF entgegenzuwirken (Kornmann 2006, Preuss-Lausitz 2001, Rauer & Schuck 2007, Wocken 2005). Allerdings bedarf es längsschnittlicher Untersuchungen, um die dargestellten Befunde abzusichern und Vermittlungsprozesse zwischen Merkmalen der Klassenzusammensetzung und den Schulleistungen von Kindern mit SPF beschreiben zu können.

Literatur

- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung. (2010). *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: W. Bertelsmann Verlag.
- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., ... Tsai, Y.-M. (2008). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Maaz, K., Stanat, P., & Watermann, R. (2009). Schulkomposition oder Institution - was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule*, 101, 33–46.
- Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, ... M. Weiß (Eds.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 323–407). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P., & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104, 682–699.
- Bell, B. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010). The impact of small cluster size on multilevel models: A Monte Carlo examination of two-level models with binary and continuous predictors (pp. 4057–4067). Presented at the JSM Proceedings, Section on Survey Research Methods, Vancouver, BC: American Statistical Association.
- Bos, W., Müller, S., & Stubbe, T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Eds.), *Bildungsverlierer. Neue Ungleichheiten* (pp. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society*. New York: John Wiley & Sons.
- Buuren van, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–68.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- De Fraine, B., Van Damme, J., Van Landeghem, G., Opdenakker, M.-C., & Onghena, P. (2003). The effect of schools and classes on language achievement. *British Educational Research Journal*, 29, 841–859.

- Diamond, J. B., Randolph, A., & Spillane, J. P. (2004). Teachers' expectations and sense of responsibility for student learning: The importance of race, class, and organizational habitus. *Anthropology & Education Quarterly*, 35, 75–98.
- Dumont, H., Neumann, M., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183.
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225–241.
- Gamoran, A. (1986). Instructional and institutional effects of ability grouping. *Sociology of Education*, 59, 185–198.
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socioeconomic index of occupational status. *Social Science Research*, 21, 1–56.
- Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention) vom 21. 12. 2008.
- Gomolla, M. (2006). Fördern und Fordern allein genügt nicht! Mechanismen institutioneller Diskriminierung von Migrantenkindern und -jugendlichen im deutschen Schulsystem. In G. Auernheimer (Ed.), *Schief lagen im Bildungssystem* (pp. 87–102). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gröhlich, C., Guill, K., Scharenberg, K., & Bos, W. (2010). Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In W. Bos & C. Gröhlich (Eds.), *KESS 8-Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (pp. 100–106). Münster: Waxmann.
- Hallinan, M. T., & Kubitschek, W. N. (1999). Curriculum differentiation and high school achievement. *Social Psychology of Education*, 3, 41–62.
- Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, 112, 1163–1197.
- Heller, K. A., & Perleth, C. (2000). *KFT 4-12+R: kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test.
- Hornstra, L., Denessen, E., Bakker, J., Bergh, L. van den, & Voeten, M. (2010). Teacher attitudes toward dyslexia: Effects on teacher expectations and the academic achievement of students with dyslexia. *Journal of Learning Disabilities*, 43, 515–529.
- Hox, J. (2002). *Multilevel analysis. Techniques and application*. London: Lawrence Erlbaum.
- Justice, L. M., Logan, J. A. R., Lin, T.-J., & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, 25, 1722–1729.
- Klein, G. (2001). Sozialer Hintergrund und Schullaufbahn von Lernbehinderten/ Förderschülern 1969 und 1997. *Zeitschrift für Heilpädagogik*, 52, 51–61.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A., & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191.
- Kocaj, A., Kuhl, P., Weirich, S., Haag, N., Pant, H. A., & Stanat, P. Zur strukturellen Validität und Messinvarianz bei der Erfassung der Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf im IQB Ländervergleich 2011. Manuskript in Vorbereitung.
- Köller, O., & Baumert, J. (2001). Leistungsgruppierungen in der Sekundarstufe I und ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie*, 15, 99–110.
- Köller, O., Schütte, K., Zimmermann, F., Retelsdorf, J., & Leucht, M. (2013). Starke Klasse, hohe Leistungen? Die Rolle der Leistungsstärke der Klasse für die individuellen Mathematik- und Leseleistungen in der Sekundarstufe I. *Psychologie in Erziehung und Unterricht*, 60, 184–197.
- Kornmann, R. (2006). Die Überrepräsentation ausländischer Kinder und Jugendlicher in Sonderschulen mit dem Schwerpunkt Lernen. In G. Auernheimer (Ed.), *Schief lagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (pp. 71–85). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Kristen, C. (2002). Hauptschule, Realschule oder Gymnasium? Ethnische Unterschiede am ersten Bildungsübergang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 54, 534–552.
- Kronig, W. (2001). Probleme der Selektion in den Grundschuljahren. *Schweizerische Zeitschrift für Bildungswissenschaften*, 23, 357–364.
- Kronig, W. (2003). Das Konstrukt des leistungsschwachen Immigrantenkindes. *Zeitschrift für Erziehungswissenschaft*, 6, 126–141.
- Lentz, A., & Radtke, F.-O. (1994). Bildungsghettos - Institutionalisierte Diskriminierung von Migrantenkindern in der Grundschule. *Unterrichtswissenschaft*, 22, 182–191.
- Maaz, K., Baumert, J., & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In J. Baumert, K. Maaz, & U. Trautwein (Eds.), *Bildungsentscheidungen* (pp. 11–46). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Markussen, E. (2004). Special education: Does it help? A study of special education in Norwegian upper secondary schools. *European Journal of Special Needs Education*, 19, 33–48.
- McElvany, N., Becker, M., & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41, 121–131.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.
- Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus*. Münster: Waxmann.
- OECD. (1999). *Classifying educational programmes. Manual for ISCED-97 implementation in OECD countries*. Paris: OECD Publishing.
- OECD. (2012). *PISA 2009 technical report*. Paris: OECD Publishing.
- Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Opdenakker, M.-C., & Damme, J. V. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, 27, 407–432.
- Opp, D. G., Budnik, I., & Fingerle, M. (2008). Sonderschulen — integrative Beschulung. In W. Helsper & J. Böhme (Eds.), *Handbuch der Schulforschung* (pp. 341–361). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Preuss-Lausitz, U. (2001). Gemeinsamer Unterricht Behinderter und Nichtbehinderter. *Zeitschrift für Erziehungswissenschaft*, 4, 209–224.
- Preuss-Lausitz, U. (2013). Inklusive Schul- und Unterrichtsentwicklung. Ein Beitrag zum Abbau von Ungleichheit. In E. Jürgens & S. Miller (Eds.), *Ungleichheit in der Gesellschaft und Ungleichheit in der Schule: Eine interdisziplinäre Sicht auf Inklusions- und Exklusionsprozesse* (pp. 171–187). Weinheim: Beltz Juventa.
- Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- Rauer, W., & Schuck, K. D. (2007). Hamburger Grundschulen und Grundschulklassen mit einer formellen Integrationsorganisation. In W. Bos, C. Gröhlich, & M. Pietsch (Eds.), *KESS 4 – Lehr- und Lernbedingungen in Hamburger Grundschulen*. Münster: Waxmann.
- Richter, D., Engelbert, M., Böhme, K., Haag, N., Hannighofer, J., Reimers, H., ... Stanat, P. (2012). Anlage und Durchführung des Ländervergleichs. In P. Stanat, H. A. Pant, K. Böhme, & D. Richter (Eds.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011* (pp. 85–102). Münster: Waxmann.
- Rist, R. C. (2000). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education (HER classic reprint). *Harvard Educational Review*, 70, 257–302.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32, 63–72.
- Ruijs, N. M., & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4(2), 67–79.

- Sälzer, C., Gebhardt, M., Müller, K., & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In: P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (pp. 129-153). Wiesbaden: Springer VS.
- Schafer, J. L., & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods*, 13, 279–313.
- Schofield, J. W. (2010). International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. *Teachers College Record*, 112, 1492–1528.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21, 43–69.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. (2nd ed.). Thousand Oaks, CA: Sage.
- Stanat, P., Pant, H. A., Böhme, K., & Richter, D. (Eds.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stocké, V. (2007). Explaining educational decision and effects of families' social class position: An empirical test of the Breen–Goldthorpe model of educational attainment. *European Sociological Review*, 23, 505–519.
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99, 253–273.
- Thrupp, M., Lauder, H., & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37, 483–504.
- Weinert, F. E., & Helmke, A. (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz.
- Wilkinson, I. A. G. (2002). Introduction: peer influences on learning: where are they? *International Journal of Educational Research*, 37, 395–401.
- Wocken, H. (2005). *Andere Länder, andere Schüler? Vergleichende Untersuchung von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen*. Retrieved from <http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf>

4.4 Teilstudie 4: Educational Placement and Achievement Motivation of Students With Special Educational Needs

Abstract

The present study investigated how the educational placement in special education schools or regular schools is related to the achievement motivation of students with special educational needs (SEN). Furthermore, we examined whether the classroom social environment (i.e., class-average achievement and social support) explains potential placement differences. We compared the academic self-concept and enjoyment of learning of SEN students in special education schools ($n = 420$) and regular schools ($n = 678$) at the end of 4th grade. In line with social comparison theory, SEN students in special education schools reported higher academic self-concepts and enjoyment of learning than their peers in regular schools. Examining underlying mechanisms, class-average achievement was negatively related to SEN students' academic self-concept and enjoyment of learning. In contrast, individually perceived social support by classmates—but not class-average social support—was positively associated with SEN students' achievement motivation. After controlling for individual and class-average achievement, no relationship between SEN students' educational placement and achievement motivation was found. Results indicate that social comparisons with classmates may result in placement differences in SEN students' academic self-concept and enjoyment of learning.

Keywords: special educational needs, educational placement, academic self-concept, enjoyment of learning, classroom composition

Draft version 06.11.2017. The review process for this manuscript has not yet been completed. Please do not copy or cite without author's permission.

Kocaj, A., Kuhl, P., Jansen, M., Pant, H. A. & Stanat, P. (2017). *Educational placement and achievement motivation of students with special educational needs*. Manuscript submitted for publication.

4.4.1 Introduction

Separate but equal is never equal (Brown v. Board of Education, 1954). But for students with special educational needs (SEN), potential benefits and drawbacks of separate schools specialized to their needs are still widely debated. With the ratification of the UN Convention on the Rights of Persons with Disabilities (United Nations, 2006), several countries have committed to implementing a more inclusive school system and to increasing the proportion of SEN students in regular schools. Although inclusive schooling involves a number of organizational features, teaching practices, and student groups, one central aspect concerns the educational placement of students with SEN. Educational placement refers to the school type (e.g., special education schools vs. regular schools) and the type of classroom where SEN students receive instruction. There is a lively debate among policy makers, school practitioners, and educational researchers (Hocutt, 1996; Lindsay, 2007): Should students with SEN be educated in regular classrooms or in separate schools or classrooms specifically targeted to their needs?

The trend towards a more inclusive school system is accompanied by intensified research on its effectiveness for SEN students (Bakker, Denessen, Bosman, Krijger, & Bouts, 2007; Lindsay, 2007; Ruijs & Peetsma, 2009). This research has not revealed consistent associations between SEN students' educational placement and aspects of their school development. With regard to cognitive outcomes, studies point to higher achievement levels of SEN students in regular schools compared to their peers in special education schools (Kocaj, Kuhl, Kroth, Pant, & Stanat, 2014; Ruijs & Peetsma, 2009). However, findings on placement differences for psychosocial (Bakker et al., 2007; Pijl & Frostad, 2010; Szumski & Karwowski, 2014) and motivational outcomes (Elbaum, 2002; Salend & Duhaney, 1999) are inconclusive or tend to favor of special education schools. Consequently, the benefits and drawbacks of inclusive schooling for SEN students seem to depend on the outcome under consideration and the attached importance of this outcome.

SEN students' achievement motivation is generally an important aspect of their school development (Elbaum, 2002; Grolnick & Ryan, 1990; Möller, Streblow, & Pohlmann, 2009). SEN students are especially at risk of developing low levels of achievement motivation because they are more likely to experience academic failure than their peers without SEN (Allodi, 2000; Grolnick & Ryan, 1990). Peetsma, Vergeer, Roeleveld, and Karsten (2001) suggested that academic failures and consequent declines in motivation could trigger a vicious

circle preventing SEN students from overcoming their learning difficulties. For example, SEN students may develop negative self-evaluations and consider their own abilities as unchangeable after experiencing repeated failure at school (Deci & Chandler, 1986; Eccles, Wigfield, & Schiefele, 1998). As a consequence, they might decrease their effort in school work and their persistence in solving demanding tasks. This, in turn, could have a negative bearing on their subsequent school achievement (Deci & Chandler, 1986). Research on students without SEN revealed that low achievement motivation is also linked to higher rates of depression, disengagement, and school dropout (Elbaum, 2002; Reyes, Brackett, Rivers, White, & Salovey, 2012). Therefore, promoting SEN students' achievement motivation is important for their success in school.

Learning environments differ between educational settings and these differences (e.g., in curricula, teaching methods, and student composition) can shape students' achievement motivation (Urdan & Schoenfelder, 2006; Wang & Eccles, 2013). Although previous studies explored placement differences in SEN students' achievement motivation (Bear, Minke, & Manning, 2002; Chapman, 1988; Elbaum, 2002; Forman, 1988; Pijl & Frostad, 2010), researchers have rarely examined characteristics of the learning environments that might underlie those differences. The purpose of this study is to investigate if SEN students' achievement motivation differs between special education schools and regular schools and to examine if the classroom social environment (i.e., social comparisons and social support) explains possible placement differences.

4.4.2 Educational Placement and SEN Students' Achievement Motivation

The educational placement of SEN students in special education schools or in regular schools could affect their achievement motivation in different ways. Proponents of special education schools argue that these settings provide a protected learning environment for SEN students. It includes adapted and less demanding curricula, individualized feedback, and a less competitive climate in the classroom (Chapman, 1988; Markussen, 2004; Peetsma et al., 2001). From a person-environment fit perspective, students should be more motivated if the learning material is adapted to their competence level (Eccles, 2004; Vaughn, Elbaum, & Schumm, 1996). Thus, the adaptation of the curricula should increase the chances that SEN students experience feelings of success. Furthermore, the achievement motivation of SEN students can

be expected to benefit from comparisons with peers with similar performance levels and similarly difficulties (Chapman, 1988). Performance comparisons with similar achieving peers should decrease SEN students' feelings of inferiority and school stress due to evaluative anxiety (Dijkstra, Kuyper, van der Werf, Buunk, & van der Zee, 2008; Renick & Harter, 1989; Tracey & Marsh, 2000).

Supporters of an inclusive education, in contrast, suggest that SEN students in regular schools could feel more appreciated as capable learners (Bakker & Bosman, 2003). This, in turn, may increase SEN students' achievement motivation (Vaughn et al., 1996; Wiener & Tardif, 2004). According to this view, teachers in regular schools have higher expectations and place stronger emphasis on academic progress (Diamond, Randolph, & Spillane, 2004; Hornstra, Denessen, Bakker, Bergh, & Voeten, 2010). By belonging to a learning group that is positively valued, SEN students in regular schools may develop higher achievement motivation through basking in reflected glory of the perceived accomplishments of their peers (Marsh et al., 2000; Tracey & Marsh, 2000). In addition, SEN students in regular schools might benefit from higher-achieving classmates through social learning mechanisms (Barth, Dunlap, Dane, Lochman, & Wells, 2004): Higher-achieving classmates could function as role models for SEN students and initiate upward comparisons associated with increases in achievement motivation (Corcoran, Crusius, & Mussweiler, 2011; Wheeler & Suls, 2005).

Placement differences might vary for different motivational factors. According to expectancy-value theory, two central antecedents of students' achievement motivation are academic self-concept and enjoyment of learning (Nagengast et al., 2011; Wigfield & Eccles, 2000). The academic self-concept is defined as students' self-evaluation in a specific academic domain or subject (Marsh, Kong, & Hau, 2000; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Enjoyment of learning reflects students' intrinsic value and subjective interest in school-related tasks. Both of these motivational variables influence school achievement and task persistence as well as educational choices and career aspirations (Eccles et al., 1983; Nagengast et al., 2011; Pekrun, Goetz, Titz, & Perry, 2002; Wang & Eccles, 2013).

In the majority of the studies, SEN students in special classes or special education schools reported higher academic self-concepts than SEN students in inclusive settings (Allodi, 2000; Bear et al., 2002; Chapman, 1988; Crabtree & Meredith, 2000; Renick & Harter, 1989; Tracey & Marsh, 2000). In a recent study by Szumski and Karwowski (2015), for example, primary students with mild intellectual disabilities in special education schools reported a

more positive academic self-concept than their peers in inclusive school settings. Bear, Minke, and Manning (2002) concluded in their meta-analysis that a more restrictive learning environment of students with learning disabilities (i.e., separated classes) is associated with higher academic self-concepts. However, they also reported considerable heterogeneity in the effect sizes across studies associated with the specific educational setting SEN students experienced. In contrast, a meta-analysis by Elbaum (2002) found no consistent relationship between educational placement and SEN students' academic self-concept. Peetsma and colleagues (2001) confirmed results from previous cross-sectional studies by applying a longitudinal design: SEN students' teacher-rated achievement motivation was significantly higher in special education schools than in regular schools. Over a period of two years, there was also a less pronounced decrease in SEN students' achievement motivation in special education schools than in regular schools. After four years, however, no placement differences in SEN students' achievement motivation were found.

Despite the large number of studies investigating placement effects on SEN students' motivational outcomes, studies have rarely explored characteristics of the classroom social environment potentially underlying those placement differences. Furthermore, the majority of the studies focused on the academic self-concept of students with learning disabilities. Although some studies included other aspects of achievement motivation (Fulk, Brigham, & Lohman, 1998; Grolnick & Ryan, 1990; Peetsma et al., 2001), research on intrinsic value components—such as enjoyment of learning—is scarce.

4.4.3 Mechanisms Underlying Placement Differences

Special educational schools and regular schools provide different social environments that could contribute to differences in SEN students' achievement motivation. Research on students without SEN has shown that achievement motivation is influenced by the classroom social environment (Patrick et al., 2007; Urdan & Schoenfelder, 2006; Wang & Eccles, 2013). Two important features of the classroom social environment are *social comparisons* and *social support* structures.

4.4.3.1 Social Comparisons as Underlying Mechanism of Placement Differences

SEN students in regular schools learn together with higher-achieving classmates without SEN whereas SEN students in special education schools learn together with other SEN students with similar difficulties. Differences between special education schools and regular schools in classmates' achievement level might result in differences in SEN students' achievement motivation. According to social comparison theory, students compare themselves with classmates to gain an accurate self-view and to adjust their self-evaluation (Festinger, 1954; Marsh et al., 2008). Drawing on this assumption, Marsh (1987) postulated the big-fish-little-pond effect (BFLPE; Marsh, 1987), suggesting that social comparisons within the external frame of reference are a central source of students' academic self-concept. The BFLPE hypothesizes that individual achievement is positively related to academic self-concept, whereas the average achievement of the reference group is negatively related to academic self-concept. Specifically, given similar individual achievement, students in classrooms (or schools) with high-achieving peers (big ponds) will show lower academic self-concepts than students in classrooms with low-achieving peers (little ponds).

This negative overall effect of class-average achievement represents the net-effect of two counterbalancing processes: assimilation and contrast. On the one hand, assimilation effects (also referred to as labeling effects or basking in reflected glory) pertain to students' tendency to form academic self-concepts that are congruent with the perceived ability level of their peers (Corcoran et al., 2011; Dijkstra et al., 2008). Attending prestigious schools or classrooms with high-achieving students—such as regular schools for SEN students—should therefore lead to more positive self-concepts. In contrast, membership in low-achieving or stigmatized learning groups—such as special education schools—should be associated with more negative self-concepts.

On the other hand, contrast effects (also referred to as reference group effect) are the result of social comparisons emphasizing dissimilarity between oneself and the local reference group (Corcoran et al., 2011). Contrast effects occur if students' academic self-concept is enhanced by lower achievement levels and decreased by higher achievement levels of classmates (Marsh et al., 2000). The low-achieving reference group in special education schools could lead SEN students to the self-evaluation that they are doing just as well academically as their classmates (Renick & Harter, 1989). This comparison with low-achieving classmates boosts their academic self-concept. SEN students in regular schools, in contrast, compare their

achievement to more able classmates. As a consequence, their academic self-concept should decrease (Marsh et al., 2000).

For students without SEN, the empirical evidence clearly suggests that negative contrast effects are stronger than positive assimilation effects, leading to an overall negative BFLPE (Marsh et al., 2000). Although most studies on the BFLPE focus on the academic self-concept, there is some evidence that the BFLPE also generalizes to other motivational outcomes including enjoyment of learning (Goetz et al., 2004), interest (Köller, Trautwein, Lüdtke, & Baumert, 2006; Schurtz, Pfof, Nagengast, & Artelt, 2014; Trautwein et al., 2006), and test anxiety (Zeidner & Schleyer, 1999). The BFLPE on other motivational outcomes has been shown to be mediated by academic self-concept (Köller et al., 2006; Trautwein et al., 2006).

4.4.3.2 Social Support as Underlying Mechanism of Placement Differences

In addition to social comparisons, social support is a second characteristic of the classroom social environment that affects students' achievement motivation (Eccles, 2004; Patrick et al., 2007; Wang & Eccles, 2013). Social support can be described as students' perception of a caring classroom in which they experience affiliation and help by teachers and classmates (Urda & Schoenfelder, 2006; Wentzel, 1998). Students' individual perception of social support is shaped by the classroom social environment which includes dyadic interactions between students and teachers, group interactions, as well as shared learning values and expectations (Connell & Wellborn, 1991; Russell, 2012). Moreover, the classroom social environment may influence students' achievement motivation over and above perceived social support on the individual level (Connell & Wellborn, 1991; Wentzel, Battle, Russell, & Looney, 2010). It is therefore important to consider both individual and class-level indicators of social support (Russell, 2012).

Social support might contribute to academic self-concept and enjoyment of learning in different ways (Forman, 1988; Wentzel et al., 2010): First, classrooms where students experience support by teachers and classmates should provide opportunities to participate in lessons without anxieties of being teased or ridiculed (Patrick, Ryan, & Kaplan, 2007). The opportunity to ask questions and reflect one's own understanding in a safe learning environment supports students' feelings of confidence and expectations for success (Furrer &

Skinner, 2003; Patrick et al., 2007). Social support thus serves as resource for maintaining positive emotions towards learning (Furrer & Skinner, 2003). Second, social support can reduce detrimental effects of stressful experiences on students' achievement motivation (Allodi, 2000; Wentzel, 1998). The perceived availability of social support in particular serves as a buffer against stressful events (Demaray & Malecki, 2002; Furrer & Skinner, 2003). Students who feel supported should be less prone to potential rejection or failure during class discussions or exams. They may also spend less time ruminating about their social standing in the classroom and instead focus their attention on the learning content (Song, Bong, Lee, & Kim, 2015).

SEN students in special education schools and regular schools may differ in their perceived social support by classmates, which in turn could contribute to differences in their achievement motivation. On the one hand, SEN students might benefit from placement in special education schools. As SEN students in special education schools are taught together with classmates with similar difficulties and achievement levels, their social standing might be less salient than in regular schools. This could lead to stronger feelings of relatedness and increase SEN students' achievement motivation. In contrast, SEN students' academic difficulties in regular schools were found to be associated with lower levels of social support by classmates (Bakker et al., 2007; Savage, 2005; Vaughn, Haager, Hogan, & Kouzekanani, 1992). In addition, deficiencies in social competencies (Bauminger, Edelsztein, & Morash, 2005) and behavioral problems (Margalit & Efrati, 1996; Vaughn et al., 1992) might be more salient for SEN students in regular schools leading to lower acceptance and social support by classmates. On the other hand, proponents of an inclusive education argue that the SEN status could have a protective function for SEN students in regular schools (Vaughn et al., 1992). Furthermore, research indicates that classmates in regular schools have higher social competencies and display lower rates of bullying than classmates in special education schools (Rose, Monda-Amaya, & Espelage, 2011) which should be associated with higher levels of social support for SEN students in regular schools.

4.4.4 The Present Study

The purpose of this study is to investigate the relationship between the educational placement of SEN students and their achievement motivation. Furthermore, we examine if the classroom social environment (i.e., class-average achievement and social support) underlies potential placement differences. Although research points to higher achievement motivation in special education schools, meta-analyses (Bear et al., 2002; Elbaum, 2002) revealed considerable variability in the magnitude of placement differences. An explanation for the variability in effect sizes could be that aspects of the social environment in special education schools and regular schools—rather than the educational placement per se—are related to SEN students' achievement motivation (Grolnick & Ryan, 1990; Ruijs, Peetsma, & van der Veen, 2010). It has been argued that research should examine specific characteristics of the learning environment that might explain placement differences in SEN students' outcomes (Lindsay, 2007; Zigmond, 2003). In this study, we address two research questions:

1. Do SEN students in special education schools and regular schools differ in their academic self-concept and enjoyment of learning?
2. Are placement differences in academic self-concept and enjoyment of learning related to differences in class-average achievement and class-average social support?

Regarding the first research question, we expected SEN students in special education schools to report higher academic self-concept and enjoyment of learning than SEN students in regular schools. Regarding the second research question, we expected that placement differences in class-average achievement and class-average social support are related to higher academic self-concept and enjoyment of learning of SEN students in special education schools. After controlling for individual achievement, SEN students in lower-achieving classrooms should report higher academic self-concept and enjoyment of learning. Furthermore, SEN students' who feel socially supported by their classmates should also report higher academic self-concept and enjoyment of learning. In addition, high class-average social support should boost SEN students' academic self-concept and enjoyment of learning. We hypothesized that the effect of educational placement would no longer be significant when differences between special education schools and regular schools in class-average achievement and class-average social support are taken into account.

4.4.5 Method

4.4.5.1 Participants and Procedure

We analyzed data from a nationally representative cross-sectional assessment study of German fourth-graders that was carried out in spring 2011 (*IQB National Assessment Study 2011*, Stanat, Pant, Böhme, & Richter, 2012). The study aimed at measuring students' levels of proficiency with standardized achievement tests based on the German national educational standards. Furthermore, questionnaires were used to gather data on students' academic self-concept and enjoyment of learning, perceived social support, and background characteristics. In this study, we examined the subsample of classrooms with at least one SEN student. The resulting student sample consisted of 8692 students from 451 classrooms (see Tabelle 20) including 1098 students with SEN (Mean age = 10.90 years, SD = 0.63; 32.7% girls; 24% students with at least one parent born abroad). Class size and number of SEN students differed between special education schools and regular schools (see Tabelle 20). In regular schools, classrooms consisted of both students with and students without SEN. In contrast, only SEN students were enrolled in special education schools.

Tabelle 20. *Sample Sizes and Descriptive Statistics by Educational Placement*

	Educational Placement	
	Special education schools	Regular schools
Student level		
<i>N</i> SEN students overall	420	678
<i>n</i> SEN students with learning disabilities	261	289
<i>n</i> SEN students with speech or language impairment	140	171
<i>n</i> SEN students with emotional disorders	19	218
Average number of SEN students per classroom (<i>SD</i>)	9.84 (4.05)	1.90 (1.36)
Range number of SEN students per classroom	2 – 26	1 – 9
<i>N</i> Students overall	420	8272
Classroom level		
<i>N</i> Classrooms	49	402
Average class size (<i>SD</i>)	9.84 (4.05)	20.58 (4.25)
Range class size	2 – 26	6 – 33

SEN = special educational needs

Students with learning disabilities ($n = 550$), speech or language impairment ($n = 311$) and emotional disorders ($n = 237$) were part of the sample (see Tabelle 20).¹² Students with learning disabilities show impairments in their learning and school achievement that can be caused by a multitude of factors reaching from problems in basic psychological processes and sensory impairment to social and emotional difficulties (KMK, 1999). Children with speech or language impairment exhibit specific language-related deficits pertaining to receptive and/or productive abilities (KMK, 1998). Students with emotional disorders experience difficulties in their regulation of emotions and their social behavior in interactions with peers and teachers (KMK, 2000). Students with functional or intellectual disabilities as well as non-native German language speakers who had attended German schools for less than one year and had very low levels of language proficiency were not tested. This is in line with the testing guidelines of other educational large-scale assessments (see Joncas & Foy, 2012; OECD, 2014).

4.4.5.2 Measures

Academic self-concept. Academic self-concept as one antecedent of achievement motivation was measured with four items separately for German and mathematics (e.g., "I am usually good at solving exercises in German [mathematics] lessons"). The scale was adopted from the TIMSS 2007 study in primary schools (Martin & Preuschoff, 2008). Responses were given on a 4-point scale ($1 = do not agree$; $4 = strongly agree$). We excluded one item because of a low corrected part-whole correlation with the scale and a small factor loading in confirmatory factor analyses. This item has a reverse phrasing which may be difficult to understand for SEN students (Marsh, 1986; Nusser, Carstensen, & Artelt, 2015). The internal consistencies of both scales were reasonably high for SEN students (German: $\alpha = .71$; mathematics: $\alpha = .79$).

¹² SEN categories overlap partially between Germany and the U.S. (Gebhardt, Sälzer, Mang, Müller, & Prenzel, 2015). SEN categories in Germany describe educational supports intended to help children with disabilities advance in their school career and master the curriculum. Conversely, SEN categories in the U.S. focus on individual disabilities (Powell, 2009). In this article, we use the names of the U.S. categories to align with the international literature. However, there are subtle differences between the German and U.S. categories which should be noted. For example, a diagnosed dyslexia or dyscalculia alone is not sufficient for the identification of SEN in Germany (Gebhardt et al., 2015).

Enjoyment of learning. Self-reported enjoyment of learning is the second antecedent of achievement motivation in our study. The scale was developed by Pekrun (1992) and consists of three items assessing students' global enjoyment in school lessons (e.g., "I enjoy acquiring new knowledge in school lessons"). Students responded on a 4-point scale ($1 = do not agree$; $4 = strongly agree$). The scale showed a satisfactory internal consistency for SEN students ($\alpha = .75$).

School achievement. Standardized achievement tests in mathematics and reading comprehension were used to assess students' levels of proficiency (for a detailed description, see Stanat et al., 2012). Test accommodations for SEN students in special education schools included a selection of easier test items (drawn from a common item pool) and shorter testing time (40 minutes instead of 80 minutes per domain). To evaluate the structural validity and comparability of the achievement tests for SEN students in special education schools and regular schools, missing patterns, Rasch model fit, differential item functioning, and correlations of the test scores were analyzed. Items used in both special education schools and regular schools had similar psychometric properties. The results suggest that test scores for both groups are comparable and can be reported on a common scale (for details, see Kocaj et al., 2016). Achievement tests for reading comprehension and mathematics were scaled using a 1-parameter logistic item response theory (IRT) model (*Rasch model*). Weighted likelihood ability estimates were used as achievement scores in the analyses (WLE, Warm, 1989).

Perceived social support. Self-reported social support was assessed with four items (e.g., "My classmates cheer me up when I am sad"). The scale focuses on emotional aspects of social support by classmates. Items were selected from an established questionnaire (FEES 3-4, Rauer & Schuck, 2003) assessing children's emotional and social school experiences in primary school. The scale showed an internal consistency of $\alpha = .63$ for SEN students.

4.4.5.3 Statistical Analyses

Prior to addressing our research questions, we checked the fit of the measurement model for each motivational outcome separately by conducting confirmatory factor analyses (CFAs). First, for SEN students in each school setting (special education school vs. regular school), a two-level measurement model with equal factor loadings on the student level (L1) and on the classroom level (L2) was specified. Second, we tested for measurement invariance for each

motivational outcome between SEN students in special education schools and regular schools. Comparable measurement models and measurement invariance are necessary requirements for valid comparisons of the latent factor means across both student groups (Van de Schoot, Lugtig, & Hox, 2012).

We conducted multilevel regression analyses in Mplus 7.0 (Muthén & Muthén, 1998-2012) to predict SEN students' academic self-concept and enjoyment of learning. Both dependent variables were modeled as latent variables (Lüdtke et al. 2008). School achievement (reading comprehension and mathematics respectively) and perceived social support were entered in the model as manifest predictors on the individual level; school type (special education school vs. regular school), class-average achievement, and class-average social support were entered as manifest predictors on the classroom level. We included individual achievement and individual perceived social support as predictors for three reasons: a) to estimate the effects of interindividual differences in both predictors on motivational outcomes, b) to account for potential differences in those predictors between SEN students in special education schools and regular schools, and c) to estimate the contextual effect of class-average achievement and class-average social support on SEN students' motivational outcomes (Marsh et al., 2008). Furthermore, gender was included as a covariate on the individual level because of its association with motivational outcomes (Frenzel, Pekrun, & Goetz, 2007; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Marsh & Yeung, 1998).

We standardized all continuous variables ($M = 0$, $SD = 1$) across the SEN student sample to facilitate the interpretation of the regression coefficients (Trautwein et al., 2006). In a first step, the aggregation of class-average achievement and class-average social support was based on all students in the classroom. That is, in regular schools, the achievement and perceived social support of students without SEN are also included in the class-average variables. In a second step, only students with SEN were analyzed in the multilevel analyses (for a similar approach, see Ruijs et al., 2010). Multilevel regression analyses were specified as random intercept models with maximum likelihood parameter estimates with robust standard errors. Missing data were handled by applying full information maximum likelihood (FIML) in Mplus.

We chose a step-wise approach of entering class-level predictors into the multilevel regression model to explain potential placement differences by differences in class-average achievement and class-average social support (for a similar approach see Trautwein et al.,

2006). First, school type (Model 1 in Tabelle 22 - 24) and class-average achievement (Model 2 in Tabelle 22 - 24) were entered separately as predictors of academic self-concept or enjoyment of learning. In Model 3 (Tabelle 22 - 24), school type and class-average achievement were entered simultaneously as predictors to test if differences in class-average achievement explain placement differences in motivational outcomes. In Model 4 (Tabelle 22 - 24), class-average social support was entered as predictor on the classroom level (instead of class-average achievement) together with school type. In Model 5 (Tabelle 22 - 24), social support and achievement were included simultaneously with school type as predictors in order to examine their specific contributions to placement differences in academic self-concept and enjoyment of learning.

4.4.6 Results

4.4.6.1 Preliminary Analyses: Model Fit and Measurement Invariance

Prior to the analyses, two CFAs were conducted for SEN students in special education schools and SEN students in regular schools for each motivational outcome separately. The measurement models showed a reasonable fit for each group (see Model 0 in Anhang B2 - B4). Next, we tested for measurement invariance (see Model 1 - 3 in Anhang B2 - B4 for the fit indices). We established partial invariance for each motivational outcome respectively by allowing the intercept of one item to differ between SEN students in special education and regular schools (see Model 1 - 3 in Anhang B2 - B4). For each motivational outcome, there were three factor loadings and at least two intercepts that were constrained equal across groups. This allows valid inferences about differences in the latent means between groups (Byrne, Shavelson & Muthén, 1989; Steinmetz, 2013). We refrain from comparing the manifest factor means, however, because full scalar equivalence could not be established (Steinmetz, 2013).

4.4.6.2 Placement Differences in Class-Average Achievement and Class-Average Social Support

To examine placement differences in the classroom social environment, we analyzed the pattern of correlations between school type, achievement, and social support (s. Tabelle 21). As expected, there were substantial negative associations between educational placement and achievement: class-average achievement was higher in regular schools than in special education schools ($r = -.83$, $p < .001$ for reading comprehension and $r = -.81$, $p < .001$ for mathematics achievement, Tabelle 21). The same pattern of results emerged for the individual level ($r = -.46$, $p < .001$ for reading comprehension and $r = -.51$, $p < .001$ for mathematics, Tabelle 21). The correlation between class-average social support and educational placement also favored regular schools but it was considerably lower than for achievement ($r = -.31$, $p < .001$). Furthermore, there was no substantial relationship between educational placement and social support on the individual level ($r = .08$, $p = .20$). These correlations indicate that SEN students in regular schools are, on average, surrounded by higher-achieving classmates that also perceive higher levels of social support than SEN students in special education schools.

Tabelle 21. *Correlations Between Educational Placement, Achievement, and Social Support*

	1	2	3
	Student level		
1 Educational Placement ^a (1 = special education school)			
2 Reading achievement (WLE)	-.46***		
3 Mathematics achievement (WLE)	-.51***	.66***	
4 Social support	.08	.05	.04
	^b Classroom level		
1 Educational Placement ^a (1 = special education school)			
2 Reading achievement (WLE)	-.83***		
3 Mathematics achievement (WLE)	-.81***	.91***	
4 Social support	-.31***	.32***	.34***

WLE = weighted likelihood ability estimates (Warm, 1989), ^aPoint-biserial correlations are reported for educational placement, ^bCorrelations on the classroom level are based on unstandardized aggregations of individual characteristics. * $p < .05$, ** $p < .01$, *** $p < .001$

4.4.6.3 *Placement Differences in Academic Self-Concept and Enjoyment of Learning and Their Relationship With Class-Average Achievement and Class-Average Social Support*

Results of the multilevel analyses are presented in Tabelle 22, 23, and 24 for academic self-concept in German and mathematics, as well as enjoyment of learning, respectively. To address our first research question, academic self-concept and enjoyment of learning were regressed on school type, individual achievement, and gender (Model 1, Tabelle 22 - 24). When controlling for individual achievement and gender, academic self-concepts in German ($b = 0.31, p < .001$) and mathematics ($b = 0.48, p < .001$) were higher for SEN students in special education schools than for their peers in regular schools. SEN students in special education schools also reported higher enjoyment of learning than SEN students in regular schools after taking individual differences in reading comprehension and gender differences into account ($b = 0.30, p < .001$).¹³ These results reflect the anticipated placement differences in favor of special education schools.

In Model 2, individual achievement and class-average achievement were entered to predict academic self-concept and enjoyment of learning (Tabelle 22 - 24). The regression coefficients of this model were in accordance with the expected BFLPE: individual achievement was positively related to academic self-concept (German: $b = 0.16, p < .001$; mathematics: $b = 0.41, p < .001$) whereas class-average achievement was negatively related to academic self-concept (German: $b = -0.20, p < .001, ES2^{14} = -0.48$; mathematics: $b = -0.35, p < .001, ES2 = -0.65$). In other words, SEN students with similar individual achievement reported lower academic self-concepts in higher-achieving classrooms. Individual achievement in reading comprehension ($b = 0.13, p = .001$) was also positively related to SEN students' enjoyment of learning (Model 2, Tabelle 24). After individual differences in achievement and gender differences were taken into account, the regression coefficient for class-average reading comprehension achievement was negative ($b = -0.18, p < .001$) —that is, the BFLPE was detected (albeit weaker, $ES2 = -0.40$) not only for academic self-concept, but also for enjoyment of learning.

¹³ Analyses with achievement in mathematics as predictor yielded similar results and are presented in Anhang B5.

¹⁴ $ES2$ is a measure of effect size in multilevel structural equation modeling and comparable to Cohen's d for continuous class-level predictors (Lüdtke et al., 2008; Marsh et al., 2009).

In Model 3, school type and characteristics of the classroom social environment were included simultaneously as predictors (Tabelle 22 - 24). The regression coefficients of class-average achievement on academic self-concept remained significantly negative (German: $b = -0.16$, $p = .008$; mathematics: $b = -0.32$, $p < .001$), whereas the regression coefficients of school type were no longer significant (German: $b = 0.09$, $p = .26$; mathematics: $b = 0.07$, $p = .34$). Similar results were obtained for enjoyment of learning (Model 3, Tabelle 24): Controlling for individual reading comprehension and gender, SEN students reported lower levels of enjoyment of learning in classrooms the higher class-average achievement in reading comprehension ($b = -0.14$, $p = .03$) was. Educational placement had no additional effect on SEN students' enjoyment of learning ($b = 0.09$, $p = .40$). These findings indicate that class-average achievement can explain placement differences in academic self-concepts and enjoyment of learning between SEN students in special education schools and regular schools.

Tabelle 22. *Predicting SEN Students' Academic Self-Concept in German*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)										
Reading achievement	0.11***	0.03	0.16***	0.03	0.15***	0.04			0.15***	0.03
Perceived social support							0.14***	0.03	0.13***	0.03
Gender (1 = female)	0.12*	0.05	0.11*	0.05	0.11*	0.05	0.09	0.05	0.09	0.05
Classroom level (L2)										
School type (1 = special education school)	0.31***	0.07			0.09	0.08	0.21***	0.06	0.07	0.08
Class-average reading achievement			−0.20 ***	0.05	−0.16**	0.06			−0.16**	0.06
Class-average social support							−0.02	0.03	−0.01	0.03
Model fit										
Chi-Square test of model fit (df)	35.94*** (9)		38.45*** (9)		39.09*** (11)		41.98*** (11)		49.48*** (17)	
Bayesian Information Criteria (BIC)	9668.56		10604.50		9665.25		9010.61		12022.62	
Akaike Information Criteria (AIC)	9578.72		10504.47		9570.41		8919.62		11908.03	
Comparative Fit Index (CFI)	0.95		0.95		0.95		0.95		0.95	
Root Mean Square Error (RMSEA)	0.05		0.06		0.05		0.06		0.04	
Described Variance										
<i>R</i> ² on student level (L1)	4.40%		7.90%		7.50%		6.10%		11.50%	
<i>R</i> ² on classroom level (L2)	26.80%		46.80%		46.60%		21.10%		46.60%	

SEN = special educational needs, Intraclass correlation (ICC) = .09, * $p < .05$, ** $p < .01$,

*** $p < .001$

In Model 4 (Tabelle 22 - 24), gender, perceived social support by classmates on the individual level (L1), and class-average social support and school type (L2) were entered as predictors of SEN students' academic self-concept and enjoyment of learning. SEN students' individual-level perceived social support was positively related to their academic self-concept in both domains (German: $b = 0.14, p < .001$; mathematics: $b = 0.12, p = .006$). However, class-average social support was not significantly related to academic self-concept (German: $b = -0.02, p = .97$; mathematics: $b = 0.06, p = .17$) after accounting for the positive effect of educational placement in special education schools (German: $b = 0.21, p < .001$; mathematics: $0.27, p < .001$). SEN students with higher levels of perceived social support also reported more enjoyment of learning ($b = 0.21, p < .001$). Furthermore, SEN students in special education schools reported more enjoyment of learning than their peers in regular schools ($b = 0.24, p = .001$). Class-average social support had no additional effect on SEN students' enjoyment of learning ($b = 0.06, p = .16$). Thus, in contrast to class-average achievement, class-average social support could not explain placement differences in SEN students' academic self-concept and enjoyment of learning.

Tabelle 23. *Predicting SEN Students' Academic Self-Concept in Mathematics*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)										
Mathematics achievement	0.30***	0.04	0.41***	0.04	0.41***	0.05			0.40***	0.04
Perceived social support							0.12**	0.04	0.10**	0.04
Gender (1 = female)	−0.13*	0.06	−0.12	0.06	−0.11	0.06	−0.26***	0.06	−0.15*	0.06
Classroom level (L2)										
School type (1 = special education school)	0.48***	0.08			0.07	0.08	0.27***	0.07	0.06	0.08
Class-average mathematics achievement			−0.35 ***	0.05	−0.32***	0.06			−0.34***	0.06
Class-average social support							0.06	0.04	0.07	0.04
Model fit										
Chi-Square test of model fit (df)	46.79*** (10)		46.34*** (10)		47.09*** (12)		28.47** (12)		62.45*** (18)	
Bayesian Information Criteria (BIC)	9595.29		10510.38		9572.08		9050.50		11941.62	
Akaike Information Criteria (AIC)	9510.61		10415.39		9482.44		8964.40		11832.16	
Comparative Fit Index (CFI)	0.95		0.95		0.96		0.98		0.95	
Root Mean Square Error (RMSEA)	0.06		0.06		0.05		0.04		0.05	
Described Variance										
<i>R</i> ² on student level (L1)	17.10%		28.10%		27.60%		5.50%		29.10%	
<i>R</i> ² on classroom level (L2)	24.60%		55.40%		55.40%		12.10%		58.30%	

SEN = special educational needs, Intraclass correlation (ICC) = .14, * $p < .05$, ** $p < .01$, *** $p < .001$

In Model 5 (Tabelle 22 - 24), social support and achievement were included simultaneously with gender and school type as predictors of academic self-concept and enjoyment of learning. On the individual level, higher levels of both perceived social support (German: $b = 0.13$, $p < .001$; mathematics: $b = 0.10$, $p = .003$) and individual achievement (German: $b = 0.15$, $p < .001$; mathematics: $b = 0.40$, $p < .001$) contributed to higher levels of academic self-concept. On the classroom level, only class-average achievement (German: $b = -0.16$, $p = .004$; mathematics: $b = -0.34$, $p < .001$) predicted SEN students' academic self-concept. Educational placement (German: $b = 0.07$, $p = .38$; mathematics: $b = 0.06$, $p = .45$) and class-average social support (German: $b = -0.01$, $p = .87$; mathematics: $b = 0.07$, $p = .07$) had no additional effect on academic self-concept. With regard to enjoyment of learning, both students with higher achievement in reading comprehension ($b = 0.12$, $p = .001$) and students with higher perceived social support ($b = 0.20$, $p < .001$) on the individual level reported

higher enjoyment of learning (Model 5, Tabelle 24). Similar to academic self-concept, class-average achievement in reading comprehension had a negative effect on students' enjoyment of learning ($b = -0.16$, $p = .02$). After controlling for class-average achievement, neither school type ($b = 0.08$, $p = .44$) nor class-average social support ($b = 0.08$, $p = .08$) were significantly associated with SEN students' enjoyment of learning (Model 5, Tabelle 24). These findings indicate that class-average achievement is decisive for explaining placement differences; class-average social support does not additionally describe differences in SEN students' achievement motivation. We also did not find differential relationships between educational placement and motivational outcomes for students with learning disabilities and students with speech or language impairment (see Anhang B6 - B8).

Tabelle 24. *Predicting SEN Students' Enjoyment of Learning With Reading Achievement*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)										
Reading achievement	0.09*	0.04	0.13**	0.04	0.13**	0.04			0.12**	0.04
Perceived social support							0.21***	0.03	0.20***	0.03
Gender (1 = female)	0.20**	0.07	0.19**	0.07	0.19**	0.07	0.14*	0.07	0.14*	0.07
Classroom level (L2)										
School type (1 = special education school)	0.30***	0.08			0.09	0.11	0.24**	0.07	0.08	0.10
Class-average reading achievement			−0.18 ***	0.05	−0.14*	0.07			−0.16*	0.07
Class-average social support							0.06	0.04	0.08	0.04
Model fit										
Chi-Square test of model fit (df)	58.40*** (9)		55.13*** (9)		54.68*** (11)		33.83*** (11)		64.55*** (17)	
Bayesian Information Criteria (BIC)	9694.95		10635.45		9696.43		8992.37		12015.39	
Akaike Information Criteria (AIC)	9605.13		10535.42		9601.62		8901.55		11900.80	
Comparative Fit Index (CFI)	0.92		0.92		0.93		0.97		0.93	
Root Mean Square Error (RMSEA)	0.07		0.07		0.06		0.05		0.05	
Described Variance										
<i>R</i> ² on student level (L1)	3.90%		5.70%		5.50%		10.90%		13.40%	
<i>R</i> ² on classroom level (L2)	12.30%		20.50%		20.00%		12.70%		24.70%	

SEN = special educational needs, Intraclass correlation (ICC) = .15, * $p < .05$, ** $p < .01$,

*** $p < .001$

4.4.7 Discussion

4.4.7.1 Summary and Interpretation

The purpose of the study was to examine how educational placement and characteristics of the classroom social environment are related to SEN students' achievement motivation. Our findings support previous research showing that SEN students in special education schools report a more positive academic self-concept and higher levels of enjoyment of learning than SEN students in regular schools (Bear et al., 2002; Chapman, 1988; Renick & Harter, 1989). Placement differences in SEN students' achievement motivation can be explained by lower class-average achievement in special education schools compared to regular schools. After taking differences in class-average achievement into account, no association between educational placement and SEN students' achievement motivation was found.

Our results are in line with social comparison theory and research on the big-fish-little-pond effect (Marsh et al., 2000). SEN students in special education schools can compare their achievement with that of low-achieving classmates which benefits their academic self-concept and enjoyment of learning. This contrast effect—being a big fish in a small pond— seems to be more important than the potential negative labeling effect of being placed in a stigmatized school type (i.e., assimilation effect). The effect sizes of the negative relationship between class-average achievement and academic self-concept after controlling for individual achievement are similar to those from samples of students without SEN (for a review, see Marsh et al., 2008).

Also, the absence of assimilation effects in our study reflects previous findings on the relation between school tracking and antecedents of achievement motivation in general (Chmielewski et al., 2013; Trautwein et al., 2006). The interpretation of the present findings, however, is limited by the substantial relationship between class-average achievement and school type which makes it difficult to separate assimilation and contrast effects as described by social comparison theory (s. Tabelle 21). A direct assessment of assimilation, i.e., asking SEN students about the perceived standing or prestige of their school would help to separate contrast and assimilation effects (Trautwein, Lüdtke, Marsh, & Nagy, 2009).

In contrast to class-average achievement, class-average social support was unrelated to SEN students' achievement motivation over and above individual-level perceptions of emotional support from peers. Thus, there was no significant contextual effect of class-average social

support that could explain placement differences in SEN students' achievement motivation. The perceived achievement of classmates seems to be the primary characteristic of the social environment in the classroom contributing to placement differences in achievement motivation. Our results indicate that class-average social support by classmates does not alleviate the negative effect of classmates' perceived achievement level on SEN students' achievement motivation (but see Allodi, 2000). On the individual level, however, higher perceived social support by peers contributed to higher academic self-concept and enjoyment of learning. This finding is consistent with results from previous studies (e.g., Forman, 1988) and might have several reasons: First, perceived social support and associated feelings of relatedness can help SEN students to deal with challenging situations in school and to maintain their motivation (Furrer & Skinner, 2003). The individually perceived availability of social support might be more relevant as a buffer against stressful events in school than the class-average support structures (Furrer & Skinner, 2003). Second, SEN students' individual perceptions of a caring classroom environment could lead to positive self-perceptions and to active participation in school lessons (Grolnick & Ryan, 1990; Pijl & Frostad, 2010).

4.4.7.2 Limitations and Future Research

Several methodological limitations of our study are worth mentioning. First, questionnaires in the present educational large-scale assessment were developed for students without SEN. In the present study, measures of academic self-concept and social support showed lower internal consistencies for students with SEN than for the student population without SEN (see Richter, Böhme, Bastian-Wurzel, Pant, & Stanat, 2014). The moderate internal consistency of the social support scale as one central predictor in our study limits the interpretation of our results. This is even more relevant for aggregated variables of individual social support. The aggregation of unreliable measures on the individual level could lead to spurious effects of the corresponding measures on the classroom level (Televantou et al., 2015). Additionally, only partial measurement invariance could be established for the motivational outcomes. Placement differences in those motivational outcomes might therefore be partly due to differences in the measurement quality between both groups.

Second, our study is limited by the cross-sectional design and it is therefore not possible to draw conclusions about the directionality of our results. We proposed that social support has an influence on SEN students' academic self-concept and enjoyment of learning. Although

students who feel rejected might develop lower academic self-concept and enjoyment of learning as a consequence of low social support, it is also reasonable to argue that students with negative self-concepts and low enjoyment of learning tend to refrain from social interactions and activities in the classroom (Pijl & Frostad, 2010). Another limitation of the cross-sectional design is that we were only able to examine between-person relationships and interpret differences between groups. Longitudinal studies are necessary to examine within-person relationships, such as whether SEN students' achievement motivation increases after experiencing social support by peers. Longitudinal designs would also allow for testing reciprocal effects between SEN students' achievement motivation and social support (Skinner & Belmont, 1993). Finally, cross-sectional studies tend to overestimate placement differences and the effects of classroom composition because researchers cannot adequately control for differences in prior motivation, achievement, and the selectivity of different school types (Baumert, Stanat, & Watermann, 2006).

In the present study, we focused on emotional social support by classmates. Future studies should include measures of social support by teachers, parents, and friends and examine their interplay with support by classmates in shaping SEN students' achievement motivation (Patrick et al., 2007; Song et al., 2015) as well as differential effects of emotional and academic components of social support (Wentzel et al., 2010). We would expect that social support by teachers, parents, and friends play a distinctive role for SEN students' achievement motivation (Song et al., 2015; Wentzel et al., 2010). Furthermore, academic support such as help and instruction to meet educational expectations might differ from emotional support in predicting SEN students' achievement motivation (Wentzel et al., 2010).

4.4.7.3 Conclusions for the Educational Placement of Students with SEN

Despite its limitations, the findings of the present study contribute to the empirical debate on educational placement of students with SEN. On the one hand, educating SEN students in special education schools together with students with similar challenges seems to be beneficial for their achievement motivation. On the other hand, the majority of the studies reports cognitive benefits for SEN students in regular schools (Ruijs & Peetsma, 2009). Previous analyses with the present data also revealed higher achievement levels for SEN students in regular schools compared to special education schools (Kocaj et al., 2014). These diverging results regarding motivational and cognitive outcomes lead to the question why motivational

benefits of special education schools are not reflected in higher achievement levels for SEN students? Or, as Marsh and Parker (1984) put it: “Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?”.

A potential explanation for these opposite placement effects might be that social comparisons are based on different motives that affect motivational and cognitive outcomes in different ways. Self-enhancement motives might trigger downward comparisons with lower-achieving classmates that result in more positive academic self-concepts (Dijkstra et al., 2008). Self-improvement motives, in contrast, might lead to upward comparisons with higher-achieving classmates and subsequently boost future performance (Dijkstra et al., 2008). Higher-achieving classmates may serve as role models and may positively affect students' persistence and effort (Gamoran, 1986; Slavin, 1996). However, the downsides of upward comparisons are lower self-evaluations and negative affective consequences (Dijkstra et al., 2008).

Overall, there seems to be a positive net effect of attending a regular school on SEN students' achievement. The positive direct effect on SEN students' achievement outweighs the negative indirect effect on SEN students' achievement motivation. These contrasting effects are also reflected in the opposed relationship of class-average achievement with SEN students' cognitive and motivational outcomes. Whereas classmates' higher achievement levels promote SEN students' individual achievement (Justice, Logan, Lin, & Kaderavek, 2014), social comparisons with more able classmates lead to lower levels of achievement motivation. However, more longitudinal studies considering different aspects of the learning environment in special education schools and regular schools are necessary to shed light on the interplay between SEN students' motivational and cognitive development (Bakker et al., 2007).

Conclusions about the optimal placement for SEN students are difficult to draw because schools vary in their implementation of inclusive education (Lindsay, 2007). However, examining the learning environment in special education schools could help to promote SEN students' achievement motivation in inclusive school settings. SEN students might benefit from a less competitive classroom environment where teachers do not emphasize social comparisons but instead rely on individually oriented judgments (Chapman, 1988, Lüdtke, Köller, Marsh, & Trautwein, 2005). Furthermore, it may be possible to foster SEN students' achievement motivation by classroom structures that provide high levels of social support (Forman, 1988). This might be especially relevant for SEN students because high levels of social support may help them to compensate for academic difficulties and to preserve a

positive self-view (Allodi, 2000; Lindsay, 2007). Inclusive education might be successful in a supportive classroom climate where contributions of every student are valued and opportunities for co-operative learning, peer tutoring, and social learning are provided.

References

- Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, 39, 1490–1528. doi: 10.1177/0149206313478188
- Allodi, M. W. (2000). Self-concept in children receiving special support at school. *European Journal of Special Needs Education*, 15, 69–78. doi: 10.1080/088562500361718
- Bakker, J. T. A., & Bosman, A. M. T. (2003). Self-image and peer acceptance of Dutch students in regular and special education. *Learning Disability Quarterly*, 26, 5–14. doi: 10.2307/1593680
- Bakker, J. T. A., Denessen, E., Bosman, A. M. T., Krijger, E.-M., & Bouts, L. (2007). Sociometric status and self-image of children with specific and general learning disabilities in Dutch general and special education classes. *Learning Disability Quarterly*, 30, 47–62. doi: 10.2307/30035515
- Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology*, 42, 115–133. doi: 10.1016/j.jsp.2003.11.004
- Baumert, J., Stanat, P., & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus [School structure and the emergence of differential learning and developmental milieus]. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bauminger, N., Edelsztein, H. S., & Morash, J. (2005). Social information processing and emotional understanding in children with LD. *Journal of Learning Disabilities*, 38, 45–61. doi: 10.1177/00222194050380010401
- Bear, G. G., Minke, K. M., & Manning, M. A. (2002). Self-concept of students with learning disabilities: A meta-analysis. *School Psychology Review*, 31, 405–427.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466. doi: 10.1037/0033-2909.105.3.456
- Chapman, J. W. (1988). Learning disabled children's self-concepts. *Review of Educational Research*, 58, 347–371. doi: 10.3102/00346543058003347
- Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, 50, 925–957. doi: 10.3102/0002831213489843
- Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in social psychology* (pp. 119–139). Oxford, UK: Wiley-Blackwell.
- Crabtree, J. W., & Meredith, C. (2000). Self-concept and social comparisons in learning disabled students attending mainstream and special schools: Does integration have an impact? In R. G. Craven & H. W. Marsh (Eds.), *Self-concept theory, research and practice: Advances for the new millennium* (pp. 187–193). Sydney: SELF Research Centre, University of Western Sydney.
- Deci, E. L., & Chandler, C. L. (1986). The importance of motivation for the future of the LD field. *Journal of Learning Disabilities*, 19, 587–594.
- Demaray, M. K., & Malecki, C. K. (2002). The relationship between perceived social support and maladjustment for students at risk. *Psychology in the Schools*, 39, 305–316. doi: 10.1002/pits.10018

- Diamond, J. B., Randolph, A., & Spillane, J. P. (2004). Teachers' expectations and sense of responsibility for student learning: The importance of race, class, and organizational habitus. *Anthropology & Education Quarterly*, 35, 75–98. doi: 10.1525/aeq.2004.35.1.75
- Dijkstra, P., Kuyper, H., van der Werf, G., Buunk, A. P., & van der Zee, Y. G. (2008). Social comparison in the classroom: A review. *Review of Educational Research*, 78, 828–879. doi: 10.3102/0034654308321210
- Eccles, J. S. (2004). Schools, academic motivation, and stage-environment. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (pp. 125–153). Hoboken, NJ: John Wiley & Sons.
- Eccles (Parsons), J., Adler, T. E., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives. Psychological and sociological approaches* (S. 75–146). San Francisco, CA: W. H. Freeman.
- Elbaum, B. (2002). The self-concept of students with learning disabilities: A meta-analysis of comparisons across different placements. *Learning Disabilities Research & Practice*, 17, 216–226. doi: 10.1111/1540-5826.00047
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. doi: 10.1177/001872675400700202
- Forman, E. A. (1988). The effects of social support and school placement on the self-concept of LD students. *Learning Disability Quarterly*, 11, 115–124. doi: 10.2307/1510989
- Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics —A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22, 497–514. doi: 10.1007/BF03173468
- Fulk, B. M., Brigham, F. J., & Lohman, D. A. (1998). Motivation and self-regulation: A comparison of students with learning and behavior problems. *Remedial and Special Education*, 19, 300–309. doi: 10.1177/074193259801900506
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95, 148–162. doi: 10.1037/0022-0663.95.1.148
- Gamoran, A. (1986). Instructional and institutional effects of ability grouping. *Sociology of Education*, 59, 185–198. doi: 10.2307/2112346
- Gebhardt, M., Sälzer, C., Mang, J., Müller, K., & Prenzel, M. (2015). Performance of students with special educational needs in Germany: Findings from Programme for International Student Assessment 2012. *Journal of Cognitive Education and Psychology*, 14, 343–356. doi: 10.1891/1945-8959.14.3.343
- Goetz, T., Pekrun, R., Zirngibl, A., Jullien, S., Kleine, M., vom Hofe, R., & Blum, W. (2004). Leistung und emotionales Erleben im Fach Mathematik. Längsschnittliche Mehrebenenanalysen [Academic achievement and emotions in mathematics: A longitudinal multilevel analysis perspective]. *Zeitschrift für Pädagogische Psychologie*, 18, 201–212. doi: 10.1024/1010-0652.18.34.201
- Grolnick, W. S., & Ryan, R. M. (1990). Self-perceptions, motivation, and adjustment in children with learning disabilities: A multiple group comparison study. *Journal of Learning Disabilities*, 23, 177–184.
- Hocutt, A. M. (1996). Effectiveness of special education: Is placement the critical factor? *The Future of Children*, 6(1), 77–102. doi: 10.2307/1602495
- Hornstra, L., Denessen, E., Bakker, J., van den Bergh, L., & Voeten, M. (2010). Teacher attitudes toward dyslexia: Effects on teacher expectations and the academic achievement of students with dyslexia. *Journal of Learning Disabilities*, 43, 515–529. doi: 10.1177/0022219409355479
- Joncas, M., & Foy, P. (2012). Sample design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011* (pp. 1–21). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Justice, L. M., Logan, J. A. R., Lin, T.-J., & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, 25, 1722–1729. doi: 10.1177/0956797614538978

- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1998). *Empfehlungen zum Förderschwerpunkt Sprache [Recommendations on special educational support in speech]*. Beschluss vom 26.06.1998. Retrieved from <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sprache.pdf>
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1999). *Empfehlungen zum Förderschwerpunkt Lernen [Recommendations on special educational support in learning]*. Beschluss vom 01.10.1999. Retrieved from <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopale.pdf>
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2000). *Empfehlungen zum Förderschwerpunkt emotionale und soziale Entwicklung [Recommendations on special educational support in emotional and social development]*. Beschluss vom 10.03.2000. Retrieved from <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/emotsozentw.pdf>
- Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A., & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf [Test quality of school achievement tests for students with special educational needs]. In V. Moser & B. Lütje-Klose (Eds.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (p. 212–234). Weinheim: Beltz Juventa.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A., & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe [Where do children with special educational needs learn better? A comparison between regular primary schools and special schools]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191. doi: 10.1007/s11577-014-0253-x
- Köller, O., Trautwein, U., Lüdtke, O., & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe [On the interplay of academic achievement, self-concept, and interest in upper secondary schools]. *Zeitschrift für Pädagogische Psychologie*, 20, 27–39. doi: 10.1024/1010-0652.20.1.27
- Lindsay, G. (2007). Educational psychology and the effectiveness of inclusive education/mainstreaming. *British Journal of Educational Psychology*, 77, 1–24. doi: 10.1348/000709906X156881
- Lüdtke, O., Köller, O., Marsh, H. W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, 30, 263–285. doi: 10.1016/j.cedpsych.2004.10.002
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13, 203–229. doi: 10.1037/a0012869
- Margalit, M., & Efrati, M. (1996). Loneliness, coherence and companionship among children with learning disorders. *Educational Psychology*, 16, 69–79. doi: 10.1080/0144341960160106
- Markussen, E. (2004). Special education: Does it help? A study of special education in Norwegian upper secondary schools. *European Journal of Special Needs Education*, 19, 33–48. doi: 10.1080/0885625032000167133
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23, 129–149. doi: 10.3102/00028312023001129
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280–295. doi: 10.1037/0022-0663.79.3.280
- Marsh, H. W., Kong, C. K., & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology*, 78, 337–349. doi: 10.1037/0022-3514.78.2.337
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, 44, 764–802. doi: 10.1080/00273170903333665

- Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, 47, 213–231. doi: 10.1037/0022-3514.47.1.213
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008). The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20, 319–350. doi: 10.1007/s10648-008-9075-6
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x
- Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, 35, 705–738. doi: 10.3102/00028312035004705
- Martin, M. O., & Preuschoff, C. (2008). Creating the TIMSS 2007 background indices. In J. F. Olson, M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS 2007 Technical Report* (pp. 281–338). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Möller, J., Streblow, L., & Pohlmann, B. (2009). Achievement and self-concept of students with learning disabilities. *Social Psychology of Education*, 12, 113–122. doi: 10.1007/s11218-008-9065-z
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.
- Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the “×” out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, 22, 1058–1066. doi: 10.1177/0956797611415540
- Nusser, L., Carstensen, C. H., & Artelt, C. (2015). Befragung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen: Ergebnisse zur Messinvarianz [Questionnaires for students with special educational needs in the area of learning: Results from multi-group analysis]. *Empirische Sonderpädagogik*, 7, 99–116.
- OECD. (2014). *PISA 2012 technical report*. Paris: OECD Publishing.
- Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, 99, 83–98. doi: 10.1037/0022-0663.99.1.83
- Peetsma, T., Vergeer, M., Roeleveld, J., & Karsten, S. (2001). Inclusion in education: Comparing pupils' development in special and regular education. *Educational Review*, 53, 125–135. doi: 10.1080/00131910125044
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91–105. doi: 10.1207/S15326985EP3702_4
- Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, 41, 359–376. doi: 10.1111/j.1464-0597.1992.tb00712.x
- Pijl, S. J., & Frostad, P. (2010). Peer acceptance and self-concept of students with disabilities in regular education. *European Journal of Special Needs Education*, 25, 93–105. doi: 10.1080/08856250903450947
- Powell, J. J. W. (2009). To segregate or to separate? Special education expansion and divergence in the United States and Germany. *Comparative Education Review*, 53, 161–187. doi: 10.1086/597816
- Rauer, W., & Schuck, K. D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen (FEES 3-4) [Questionnaire assessing emotional and social experiences of third and fourth graders]*. Göttingen: Beltz Test.
- Renick, M. J., & Harter, S. (1989). Impact of social comparisons on the developing self-perceptions of learning disabled students. *Journal of Educational Psychology*, 81, 631–638. doi: 10.1037/0022-0663.81.4.631

- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104, 700–712. doi: 10.1037/a0027268
- Richter, D., Böhme, K., Bastian-Wurzel, J., Pant, H. A., & Stanat, P. (2014). *IQB-Ländervergleich 2011. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente [IQB National Assessment Study 2011. Documentation of survey instruments]*. Berlin: Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen.
- Rose, C. A., Monda-Amaya, L. E., & Espelage, D. L. (2011). Bullying perpetration and victimization in special education: A review of the literature. *Remedial and Special Education*, 32, 114–130. doi: 10.1177/0741932510361247
- Ruijs, N. M., & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4, 67–79. doi: 10.1016/j.edurev.2009.02.002
- Ruijs, N., Peetsma, T., & van der Veen, I. (2010). The presence of several students with special educational needs in inclusive education and the functioning of students with special educational needs. *Educational Review*, 62, 1–37. doi: 10.1080/00131910903469551
- Russell, S. L. (2012). *Individual- and classroom-level social support and classroom behavior in middle school* (Doctoral dissertation). Retrieved from <http://drum.lib.umd.edu/handle/1903/13204>
- Salend, S. J., & Duhaney, L. M. G. (1999). The impact of inclusion on students with and without disabilities and their educators. *Remedial and Special Education*, 20, 114–126. doi:10.1177/074193259902000209
- Savage, R. (2005). Friendship and bullying patterns in children attending a language base in a mainstream school. *Educational Psychology in Practice*, 21, 23–36.
- Schurtz, I. M., Pfof, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, 34, 32–41. doi: 10.1016/j.learninstruc.2014.08.001
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571–581. doi: 10.1037/0022-0663.85.4.571
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21, 43–69. doi: 10.1006/ceps.1996.0004
- Song, J., Bong, M., Lee, K., & Kim, S. (2015). Longitudinal investigation into the role of perceived social support in adolescents' academic motivation and achievement. *Journal of Educational Psychology*, 107, 821–841. doi: 10.1037/edu0000016
- Stanat, P., Pant, H. A., Böhme, K., & Richter, D. (Eds.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011 [Student competencies in German and mathematics at the end of the 4th grade: Results of the IQB National Assessment Study 2011]*. Münster: Waxmann.
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 9, 1–12. doi: 10.1027/1614-2241/a000049
- Szumski, G., & Karwowski, M. (2014). Psychosocial functioning and school achievement of children with mild intellectual disability in Polish special, integrative, and mainstream schools. *Journal of Policy and Practice in Intellectual Disabilities*, 11, 99–108. doi: 10.1111/jppi.12076
- Szumski, G., & Karwowski, M. (2015). Emotional and social integration and the big-fish-little-pond effect among students with and without disabilities. *Learning and Individual Differences*, 43, 63–74. doi: 10.1016/j.lindif.2015.08.037
- Televantou, I., Marsh, H. W., Kyriakides, L., Nagengast, B., Fletcher, J., & Malmberg, L.-E. (2015). Phantom effects in school composition research: Consequences of failure to control biases due to measurement error in traditional multilevel models. *School Effectiveness and School Improvement*, 26, 75–101. doi: 10.1080/09243453.2013.871302
- Tracey, D. K., & Marsh, H. W. (2000). Self-concepts of primary students with mild intellectual disabilities: Issues of measurement and educational placement. In R. G. Craven & H. W. Marsh

- (Eds.), *Self-concept theory, research and practice: Advances for the new millennium* (pp. 419–425). Sydney: SELF Research Centre, University of Western Sydney.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98*, 788–806. doi: 10.1037/0022-0663.98.4.788
- Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology, 101*, 853–866. doi: 10.1037/a0016306
- United Nations General Assembly (UN). Convention on the Rights of Persons with Disabilities, Pub. L. No. A/RES/61/106, Annex I (2006). Retrieved from <http://www.refworld.org/docid/4680cd212.html>
- Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44*, 331–349. doi: 10.1016/j.jsp.2006.04.003
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*, 486–492. doi: 10.1080/17405629.2012.686740
- Vaughn, S., Elbaum, B. E., & Schumm, J. S. (1996). The effects of inclusion on the social functioning of students with learning disabilities. *Journal of Learning Disabilities, 29*, 598–608.
- Vaughn, S., Haager, D., Hogan, A., & Kouzekanani, K. (1992). Self-concept and peer acceptance in students with learning disabilities: A four- to five-year prospective study. *Journal of Educational Psychology, 84*, 43–50. doi: 10.1037/0022-0663.84.1.43
- Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*, 12–23. doi: 10.1016/j.learninstruc.2013.04.002
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427–450.
- Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*, 202–209. doi: 10.1037/0022-0663.90.2.202
- Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology, 35*, 193–202. doi: 10.1016/j.cedpsych.2010.03.002
- Wheeler, L., & Suls, J. (2005). Social comparison and self-evaluations of competence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 566–578). New York, NY, US: Guilford Publications.
- Wiener, J., & Tardif, C. Y. (2004). Social and emotional functioning of children with learning disabilities: Does special education placement make a difference? *Learning Disabilities Research & Practice, 19*, 20–32. doi: 10.1111/j.1540-5826.2004.00086.x
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. doi: 10.1006/ceps.1999.1015
- Zeidner, M., & Schleyer, E. J. (1999). The big-fish–little-pond effect for academic self-concept, test anxiety, and school grades in gifted children. *Contemporary Educational Psychology, 24*, 305–329. doi: 10.1006/ceps.1998.0985
- Zigmond, N. (2003). Where should students with disabilities receive special education services? Is one place better than another? *The Journal of Special Education, 37*, 193–199. doi: 10.1177/0022466903037003090

5

Diskussion

5. Diskussion

Die Frage, ob Schülerinnen und Schüler mit SPF in separaten Förderschulen oder in allgemeinen Schulen gemeinsam mit Schülerinnen und Schülern ohne SPF unterrichtet werden sollten, wird im Zuge der Inklusionsbemühungen im deutschen Bildungssystem verstärkt diskutiert (Autorengruppe Bildungsberichterstattung, 2014, 2016; Grosche, 2015). Der Ausbau des gemeinsamen Unterrichts ist eine der umfangreichsten Reformen im deutschen Bildungswesen der letzten 20 Jahren (Baumert, 2016a; Weishaupt, 2016). Dieser Ausbau bedarf einer empirischen Überprüfung der Zusammenhänge und Auswirkungen einer gemeinsamen Beschulung von Schülerinnen und Schülern mit SPF und Schülerinnen und Schülern ohne SPF (UN-BRK, 2008; Wrase, 2015). In der vorliegenden Dissertation wurden Zusammenhänge zwischen der Beschulungsart und den schulischen Outcomes von Kindern mit SPF in vier Teilstudien untersucht.

In der folgenden Diskussion werden zunächst zentrale Befunde der vier Teilstudien zusammengefasst und diskutiert (Kapitel 5.1). Daran schließt sich eine Beschreibung methodischer Herausforderungen bei der Bestimmung von Schularteffekten an (Kapitel 5.2). In Kapitel 5.3 werden dann die Zusammenhänge zwischen Beschulungsart und schulischen Outcomes von Schülerinnen und Schülern mit SPF diskutiert. Anschließend werden Implikationen der vorliegenden Befunde für das Forschungsfeld aufgezeigt (Kapitel 5.4). Die Diskussion endet mit einem Ausblick zu möglichen Implikationen der Befunde für die pädagogische Praxis (Kapitel 5.5) und einem Fazit (Kapitel 5.6).

5.1 Zusammenfassung und Interpretation der Befunde

5.1.1 Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (Teilstudie 1)

Bevor die schulischen Outcomes von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen miteinander verglichen werden, sollte zunächst geprüft werden, ob die eingesetzten Testinstrumente für einen solchen Vergleich geeignet sind. In der ersten Teilstudie wurde deshalb untersucht, ob die im IQB-Ländervergleich 2011 eingesetzten Kompetenztests ähnliche Messeigenschaften und Testgütekriterien für Kinder mit SPF in

Förder- und Grundschulen aufweisen wie für die Schülergruppe ohne SPF. Das ist eine Voraussetzung für valide Schulartvergleiche (Lane & Leventhal, 2015, Tiffin-Richards & Pant, 2017; s. Kapitel 2.4).

Die Ergebnisse der ersten Teilstudie weisen darauf hin, dass die Testwerte von Kindern mit SPF zwischen Förder- und Grundschulen vergleichbar interpretiert werden können. Die Kompetenztests wiesen bei Kindern mit SPF in Förderschulen und Grundschulen zufriedenstellende Reliabilitäten und ähnliche Passungen des Raschmodells auf. Außerdem konnte für die überwiegende Mehrheit der Testitems keine stark differenziellen Itemfunktionen (DIF) festgestellt werden. Darüber hinaus zeigten sich für beide Schülergruppen mit SPF überwiegend ähnliche Zusammenhänge zwischen den Testwerten unterschiedlicher Kompetenzbereiche untereinander und den kognitiven Grundfähigkeiten. Somit können die schulischen Kompetenzen im Lesen, Zuhören und in Mathematik der Kinder mit SPF in Förder- und Grundschulen miteinander verglichen werden, ohne dass Schulartvergleiche durch differenzielle Messeigenschaften der Kompetenztests stark verzerrt werden.

Allerdings wird die vergleichbare Interpretation der Testwerte von Kindern mit SPF in Förder- und Grundschulen dadurch eingeschränkt, dass Kinder mit SPF in Förderschulen einen höheren Anteil der vorgelegten Testitems nicht bearbeiteten als Kinder mit SPF in allgemeinen Schulen. Insbesondere erreichten Förderschülerinnen und -schüler seltener das Ende der Kompetenztests. Diese Ergebnisse weisen darauf hin, dass die eingesetzten Testhefte für Kinder mit SPF in Förderschulen zu umfangreich sind, um sie in der vorgegebenen Zeit vollständig bearbeiten zu können. Jedoch wäre die Vergleichbarkeit der Testwerte zwischen den Schülergruppen durch differenzielle Anteile fehlender Werte nur dann eingeschränkt, wenn diese mit konstruktfernen Merkmalen zusammenhängen. Dazu gehören zum Beispiel Erfahrungen mit Schulleistungstests oder den in Kompetenztests eingesetzten Antwortformaten (Lane & Leventhal, 2015). In zukünftigen Untersuchungen sollte daher geprüft werden, ob fehlende Antworten durch unterschiedliche Erfahrungen mit Kompetenztests und den verwendeten Antwortformaten zwischen Schülerinnen und Schülern mit SPF in Förder- und Grundschulen erklärt werden können. Das würde dafür sprechen, dass diese Schülerinnen und Schüler im Rahmen der Kompetenztests ausführlichere Instruktionen und Beispielaufgaben benötigen und die Testbearbeitung vorher geübt werden sollte (Pohl et al., 2016).

Darüber hinaus waren die Aufgaben für Kinder mit SPF in Förderschulen (und z. T. auch für Kinder mit SPF in Grundschulen) deutlich schwieriger als für Kinder ohne SPF. Die ausgeprägten Itemschwierigkeiten für Förderschülerinnen und -schüler tragen dazu bei, dass die

Items Fähigkeitsunterschiede innerhalb dieser Schülergruppe nicht ausreichend differenzieren können (Abedi et al., 2011; Südkamp, Pohl, Hardt et al., 2015). Außerdem kann die Messgenauigkeit der Kompetenztests für Kinder mit SPF verringert sein, wenn sie in ihrer Schwierigkeit deutlich über dem zu erwartenden Fähigkeitsniveau der Schülergruppe mit SPF liegen. So weisen Kompetenztests, die mit Hilfe von Item-Response-Modellen ausgewertet werden, im mittleren Fähigkeitsbereich der gesamten Schülerpopulation die höchste Reliabilität und die geringsten Standardfehler auf (Embretson & Reise, 2000; Lane & Leventhal, 2015).

Diesen methodischen Herausforderungen kann durch die Entwicklung und Erprobung von speziellen Testaufgaben für Kinder und Jugendliche mit SPF begegnet werden (s. Kapitel 2.4.3). So wurden im Rahmen des IQB-Bildungstrends 2016 einfachere Testaufgaben entwickelt und erprobt, um eine höhere Passung zwischen Itemschwierigkeiten und Fähigkeiten der Schülergruppe mit SPF zu erzielen (Böhme, Kohrt, Haag & Weirich, 2017). Neben einer besseren Differenzierung der Fähigkeiten könnte der Einsatz leichter Aufgaben auch die Testmotivation der Schülerinnen und Schüler mit SPF erhöhen, da Itemschwierigkeiten und Fähigkeitsniveau dadurch besser aufeinander abgestimmt werden. So berichten insbesondere Schülerinnen und Schüler am unteren Fähigkeitsspektrum eine höhere Anstrengungsbereitschaft, wenn Testaufgaben an ihr Fähigkeitsniveau angepasst werden (Asseburg & Frey, 2013).

Werden leichtere Aufgaben zur Erhöhung der Präzision der Messung schulischer Kompetenzen von Schülerinnen und Schülern mit SPF eingesetzt, kann das aber auch zu Herausforderungen führen. Zum einen kann dadurch die vergleichbare Interpretation der Testwerte zur Schülergruppe ohne SPF eingeschränkt sein (Kettler & Elliott, 2010). Es muss sichergestellt werden, dass die vereinfachten Aufgaben dasselbe Konstrukt erfassen wie die regulären Aufgaben. Diese Frage der Konstruktäquivalenz bezieht sich sowohl auf inhaltliche Aspekte, die zum Beispiel mit Expertenurteilen über die Aufgaben erfasst werden können (Tiffin-Richards & Pant, 2017), als auch auf Messeigenschaften. Eine weitere Herausforderung besteht darin, zu entscheiden, welche Schülerinnen und Schüler leichtere Testaufgaben und gegebenenfalls weitere Akkommodationen benötigen (Elliott et al., 2011). Die Schülergruppe mit SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* lässt sich in manchen Fällen nur schwer von leistungsschwachen Schülerinnen und Schülern oder Kindern und Jugendlichen mit anderen Förderbedarfen (z. B. bei Teilleistungsstörungen) abgrenzen (Eberwein, 1996). Außerdem könnten Kinder und Jugendliche mit SPF in Abhän-

gigkeit ihres Förderschwerpunktes unterschiedliche Anpassungen bei der Kompetenztestung benötigen (Lane & Leventhal, 2015; Tindal & Anderson, 2011).

In Large-Scale-Assessments im Bildungsbereich steht man zudem vor der logistischen Herausforderung, leichtere Testaufgaben und auch andere Akkommodationen unter standardisierten Testbedingungen zu implementieren (Elliott et al., 2011). Hier könnten adaptive Testverfahren einen Lösungsansatz darstellen, um eine höhere Passung zwischen Itemschwierigkeiten und Personenfähigkeiten herzustellen und individuell ausgerichtete Akkommodationen einzusetzen (Thurlow, Lazarus, Albus & Hodgson, 2010). Dabei bedarf es aber zusätzlicher Forschung zur Frage, wie sich Akkommodationen in Interaktion mit weiteren Merkmalen der Testgestaltung und Schülermerkmalen auf die Testbearbeitung auswirken (Elliott et al., 2011; Tindal & Anderson, 2011).

Zusammenfassend lässt sich trotz der genannten Einschränkungen ein positives Fazit ziehen: die schulischen Kompetenzen von Kindern mit SPF in den Bereichen *Lernen, Sprache und emotionale und soziale Entwicklung* im Primarbereich können reliabel und valide erfasst werden und lassen sich auf einer gemeinsamen Skala mit der Schülergruppe ohne SPF abbilden. Diese Schlussfolgerung wird auch in internationalen Studien zur Kompetenztestung von Schülerinnen und Schülern mit SPF gezogen (Cook et al., 2009; Huynh et al., 2004; Kato et al., 2009). Die Ergebnisse der ersten Teilstudie ergänzen den deutschsprachigen Forschungsstand um eine differenzierte Betrachtung zur Vergleichbarkeit der Testwerte innerhalb der Schülergruppe mit SPF. In den bisherigen Studien zu Aspekten der Testgüte von Schulleistungstests für Kinder und Jugendliche mit SPF wurden Schülerinnen und Schüler mit SPF in Förderschulen mit Schülerinnen und Schülern ohne SPF in allgemeinen Schulen verglichen und fanden zum Teil abweichende Ergebnisse (Müller et al., 2014; Pohl et al., 2016; Südkamp, Pohl, Hardt et al., 2015).

5.1.2 Vergleich der schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen (Teilstudie 2)

Nachdem in der ersten Teilstudie geprüft wurde, ob die im IQB-Ländervergleich 2011 eingesetzten Kompetenztests für Kinder mit SPF geeignet sind, wurde in der zweiten Teilstudie der Zusammenhang zwischen Beschulungsart und schulischen Kompetenzen von Kindern mit SPF am Ende der vierten Jahrgangsstufe untersucht. Um Unterschiede zwischen Kindern

mit SPF in Förderschulen und allgemeinen Schulen in Merkmalen zu kontrollieren, die sowohl für die Zuweisung zu einer Schulart als auch für die schulischen Kompetenzen relevant sind, wurden Propensity-Score-Matching-Verfahren eingesetzt (Ho, Imai, King & Stuart, 2007; Rosenbaum & Rubin, 1983, 1985).

Im Ergebnis zeigte sich, dass Kinder mit SPF in Grundschulen höhere Kompetenzen im Lesen, Zuhören und in Mathematik aufwiesen als vergleichbare Kinder mit SPF in Förderschulen. Diese Unterschiede entsprachen einem Vorsprung zugunsten der Kinder mit SPF in Grundschulen von knapp einem halben Schuljahr im Lesen (Behrens, Böhme & Krelle, 2009; Bremerich-Vos & Böhme, 2009), annähernd einem Schuljahr im Zuhören (Behrens et al., 2009; Bremerich-Vos & Böhme, 2009) und knapp über einem halben Schuljahr in Mathematik (Reiss & Winkelmann, 2009). Insbesondere Kinder mit dem Förderschwerpunkt *Lernen* profitierten vom gemeinsamen Unterricht in Grundschulen. Ihr Leistungsvorsprung im Vergleich zu Kindern mit ähnlichen Lernvoraussetzungen in Förderschulen lag bei annähernd einem Schuljahr im Lesen, über einem Schuljahr im Zuhören und einem Schuljahr in Mathematik. Im Gegensatz dazu waren die Leistungsunterschiede zwischen Förderschulen und Grundschulen für Kinder mit dem Förderschwerpunkt *Sprache* weniger deutlich. Im Kompetenzbereich Lesen betrug der Unterschied etwas unter einem halben Schuljahr und im Zuhören annähernd einem Schuljahr zugunsten der Kinder in Grundschulen. In Mathematik konnte hingegen kein bedeutsamer Leistungsunterschied für Kinder mit dem Förderschwerpunkt *Sprache* in Abhängigkeit ihrer Beschulungsart nachgewiesen werden.

Mögliche Erklärungen für diese förderschwerpunktspezifischen Ergebnisse könnten in unterschiedlichen Lernvoraussetzungen der beiden Schülergruppen und in unterschiedlichen Lehrplänen für die beiden Förderschwerpunkte liegen. So werden bei Kindern und Jugendlichen mit dem Förderschwerpunkt *Lernen* überdauernde und bereichsübergreifende Lernschwierigkeiten festgestellt (KMK, 1999; Lauth et al., 2014). Bei Schülerinnen und Schülern mit dem Förderschwerpunkt *Sprache* beziehen sich Lernschwierigkeiten hingegen stärker auf den produktiven und rezeptiven Sprachgebrauch (KMK, 1998). Darüber hinaus verfügen Schülerinnen und Schüler mit dem Förderschwerpunkt *Sprache* über günstigere Lernvoraussetzungen (z. B. höheres bereichsspezifisches Vorwissen) als Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* (Hanushek, Kain & Rivkin, 2002; Morgan et al., 2011; Wei, Blackorby & Schiller, 2011). Außerdem orientieren sich die Lehrpläne in Förderschulen mit dem Förderschwerpunkt *Sprache* insbesondere in Mathematik stärker an denen der allgemeinen Schule (KMK, 1998) als die Lehrpläne in Förderschulen mit dem Förderschwerpunkt

Lernen. Gemeinsam könnten diese Unterschiede in den Lernvoraussetzungen und den Lehrplänen dazu beitragen, dass sich die Schularteffekte zwischen Schülerinnen und Schülern mit den Förderschwerpunkten *Lernen* und *Sprache* unterscheiden. Allerdings müssten diese Erklärungsansätze in zukünftigen Studien explizit geprüft werden.

Die Resultate der zweiten Teilstudie stimmen mit dem nationalen und internationalen Forschungsstand zum Zusammenhang zwischen Beschulungsart und schulischen Kompetenzen von Schülerinnen und Schülern mit SPF überein (Haeberlin, 1991; Lindsay, 2007; Ruijs & Peetsma, 2009; Wocken, 2000, 2005; Werning & Reiser, 2008). Vorangegangene Studien weisen mehrheitlich auf Leistungsvorteile von Kindern und Jugendlichen mit SPF in allgemeinen Schulen im Vergleich zu Förderschulen beziehungsweise separaten Förderklassen hin. In längsschnittlichen Studien waren diese Leistungsvorsprünge allerdings nicht so deutlich ausgeprägt wie in der vorliegenden Studie (Stranghöner et al., 2017). Eine Erklärung dafür ist, dass unterschiedliche Lernvoraussetzungen und Selektionseffekte in längsschnittlichen Studien besser berücksichtigt werden können als in querschnittlichen Studien (Baumert et al., 2006). So wurden die deutlichen Ausgangsunterschiede in den schulischen Kompetenzen zugunsten von Kindern mit SPF in allgemeinen Schulen bei Stranghöner et al. (2017) berücksichtigt und stärker auf die Leistungsentwicklung der Schülerinnen und Schüler mit SPF fokussiert. Im Gegensatz dazu stand in dieser Dissertation kein adäquates Maß für die Ausgangsleistungen der Schülerinnen und Schüler zur Verfügung.

5.1.3 Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf (Teilstudie 3)

In Teilstudie 2 zeigte sich, dass Kinder mit SPF in Grundschulen höhere schulische Kompetenzen erreichen als vergleichbare Kinder mit SPF in Förderschulen. Leistungsunterschiede von Kindern und Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen könnten dabei auf mindestens drei Mechanismen zurückzuführen sein: (1) auf unterschiedliche Lernvoraussetzungen der Kinder mit SPF in Förder- und Grundschulen (Morgan et al. 2008; Stanovich, 1986), (2) auf schulartspezifische Lehrpläne, didaktische Traditionen und Merkmale der Lehrkräfte und der Unterrichtsorganisation (*institutionelle Effekte*; Baumert et al. 2006; Gamoran, 1986) und (3) auf schulartspezifische Schülerzusammensetzungen (*kompositionelle Effekte*; Dreeben & Barr, 1988; Wilkinson, 2002). In Teilstudie 3 wurde die fähig-

keitsbezogene und soziale Schülerzusammensetzung als eine mögliche Erklärung für Schular-
tunterschiede betrachtet.

Die Ergebnisse von Teilstudie 3 zeigen, dass das mittlere Fähigkeitsniveau der Schulklassen positiv mit den schulischen Kompetenzen von Kindern mit SPF in Mathematik und im Leseverständnis zusammenhängen. Kinder mit SPF erzielten in leistungstärkeren Klassen höhere schulische Kompetenzen als Kinder in Klassen mit einem geringeren Fähigkeitsniveau. Die soziale Klassenzusammensetzung hatte darüber hinaus keine zusätzliche Vorhersagekraft für die Leistungsunterschiede zwischen Kindern mit SPF in Förder- und Grundschulen. Unter Berücksichtigung der fähigkeitsbezogenen Klassenzusammensetzung verringerten sich die Zusammenhänge zwischen Beschulungsart und schulischen Kompetenzen. Dies spricht dafür, dass kompositionelle Effekte neben institutionellen Effekten zur Erklärung von Leistungsunterschieden von Kindern mit SPF in Förder- und Grundschulen beitragen.

Kinder und Jugendliche mit SPF könnten dabei von Unterricht in leistungstärkeren Klassenverbänden durch die Anwesenheit positiver Lernvorbilder und die Übernahme von Lösungswegen der Mitschülerinnen und Mitschüler profitieren (Brodesser, 2016; Scharenberg, 2012; Slavin, 1996; Tillmann & Wischer, 2006). Weiterhin kann in leistungstärkeren Schulklassen ein positives Lernklima entstehen, welches sich auf die Lernmotivation und Unterrichtsgestaltung auswirkt und dadurch zu höheren schulischen Leistungen führt (Agirdag et al., 2012; Eccles & Roeser, 2011). Zum anderen können kompositionelle Effekte durch Interaktionen zwischen Schülerinnen und Schülern und Lehrkräften vermittelt werden. So passen die Lehrkräfte ihre Unterrichtsgestaltung und Erfolgserwartungen an die Klassenzusammensetzung an (Diamond et al., 2004; Rist, 2000). Die Übernahme solcher Erwartungen durch die Schülerinnen und Schüler könnte im Sinne einer sich selbsterfüllenden Prophezeiung einen Einfluss auf deren Schulleistungen ausüben (Hornstra et al., 2010; Ready & Wright, 2011).

Im deutschen Bildungssystem sind institutionelle und kompositionelle Effekte nur schwer voneinander trennbar, da Schülerinnen und Schüler vor allem aufgrund ihrer schulischen Leistungen verschiedenen Schularten zugewiesen werden (Baumert, Maaz, Stanat & Watermann, 2009; Trautwein, Lüdtke, Marsh & Nagy, 2009). Durch diese Zuweisung unterscheidet sich die fähigkeits- beziehungsweise leistungsbezogene Klassenkomposition zwischen den Schularten (Dumont et al., 2013; Köller & Baumert, 2001). Zudem können institutionelle und kompositionelle Effekte auf die schulischen Leistungen interagieren. So scheint die Zuweisung von Kindern und Jugendlichen zu unterschiedlichen Schularten insbesondere dann zur Verstärkung von Leistungsunterschieden zu führen, wenn sie mit einer curricularen Differen-

zierung einhergeht (Schofield, 2010). In zukünftigen Studien sollten diese möglichen Interaktionen näher geprüft werden (Baumert et al., 2006). Zudem sollten potenzielle Vermittlungswege überprüft werden, die dem positiven Zusammenhang zwischen mittlerem Fähigkeitsniveau der Klasse und individuellen Kompetenzständen zugrunde liegen (Wilkinson, 2002).

5.1.4 Vergleich der schulischen Motivation von Kindern mit sonderpädagogischem Förderbedarf in Förder- und Grundschulen (Teilstudie 4)

Neben der Förderung schulischer Kompetenzen ist die Stärkung der schulischen Motivation ein weiteres Ziel schulischer Bildung (Deci & Chandler, 1986; Eccles & Wigfield, 2002; Hofer, 2014). Insbesondere bei der Evaluation inklusiver Bemühungen ist es wichtig, sowohl kognitive als auch psychosoziale und motivationale Merkmale der schulischen Entwicklung im Kindes- und Jugendalter zu berücksichtigen (Prengel, 2013). Deshalb wurde in der vierten Teilstudie untersucht, ob sich Kinder mit SPF in Förder- und Grundschulen in ihrer schulischen Motivation unterscheiden. Als Merkmale der schulischen Motivation wurden das fachspezifische akademische Selbstkonzept und die fachübergreifende Lernfreude betrachtet. Dabei zeigten sich motivationale Vorteile zugunsten einer Beschulung von Schülerinnen und Schülern mit SPF in Förderschulen. Kinder mit SPF in Förderschulen berichteten ein höheres akademisches Selbstkonzept in Mathematik und in Deutsch sowie mehr Lernfreude als Kinder mit SPF in Grundschulen. Diese Schulartunterschiede ließen sich auf die mittlere Klassenleistung in Förder- und Grundschulen zurückführen. Unter Berücksichtigung der individuellen Schulleistungen berichteten Kinder mit SPF in leistungsstärkeren Klassen eine geringere schulische Motivation als Kinder mit SPF in leistungsschwächeren Klassen.

Die Ergebnisse stehen in Einklang mit der Mehrzahl vorangegangener Studien zum Zusammenhang zwischen Beschulungsart und schulischer Motivation von Schülerinnen und Schülern mit SPF (Bear et al., 2002; Haeberlin, 1991; Rheinberg & Entstrup, 1977; Tent et al., 1991; Tracey & Marsh, 2000). Allerdings wiesen einzelne aktuelle Studien keine bedeutsamen motivationalen Unterschiede zwischen Kindern und Jugendlichen mit SPF in Förderschulen und allgemeinen Schulen aus (Rossmann et al., 2011; Wild et al., 2015). Mögliche Erklärungen für diese fehlende Korrespondenz in den Ergebnissen könnten die Stichprobengröße und Unterschiede in der Auswertungsmethodik sein. Im Gegensatz zu Teilstudie 4 wurden in den zitierten Studien (Rossmann et al., 2011; Wild et al., 2015) ausschließlich

Schülerinnen und Schüler mit dem Förderschwerpunkt *Lernen* betrachtet, wobei die Stichprobengrößen insgesamt geringer waren als in Teilstudie 4. Außerdem wurden weder individuelle Leistungsunterschiede noch Unterschiede in der mittleren Klassenleistung zwischen Förderschulen und allgemeinen Schulen in den Analysen berücksichtigt.

Die Ergebnisse der vierten Teilstudie erweitern den internationalen und nationalen Forschungsstand dahingehend, dass neben individuellen Leistungsunterschieden und der Beschulungsart auch die mittlere Klassenleistung in die Analysen einbezogen wurde. Dadurch war es möglich, Fischteicheffekte (Big-Fish-Little-Pond-Effect; Marsh, 1987) für Schülerinnen und Schüler mit SPF zu testen. Dabei zeigten sich deutliche Fischteicheffekte (s. Kapitel 2.5.4.2): Kinder mit SPF berichteten in leistungsschwächeren Bezugsgruppen in Förderschulen ein höheres akademisches Selbstkonzept und mehr Lernfreude als Kinder mit SPF mit ähnlichen individuellen Leistungen in leistungstärkeren Klassenverbänden in allgemeinen Schulen. Der Zusammenhang zwischen mittlerer Leistungsstärke der Klasse und dem individuellen akademischen Selbstkonzept von Kindern mit SPF war dabei in ihrer Stärke vergleichbar zu diesem Zusammenhang innerhalb der Schülergruppe ohne SPF (Marsh et al., 2008).

Die vorliegenden Ergebnisse können dahingehend interpretiert werden, dass Schülerinnen und Schüler mit SPF die Leistungsstärke ihrer sozialen Vergleichsgruppe bei der Ausbildung ihres akademischen Selbstkonzepts und ihrer Lernfreude berücksichtigen. In Förderschulen werden Schülerinnen und Schüler mit SPF gemeinsam mit Mitschülerinnen und Mitschülern mit ähnlichen Lernschwierigkeiten unterrichtet; dadurch entspricht ihre Vergleichsgruppe stärker dem eigenen Fähigkeitsniveau. Der Vergleich mit dieser Gruppe fördert die Ausbildung positiver Selbsteinschätzungen und kann zu einer höheren Lernfreude beitragen (Kontrasteffekte; Marsh et al., 2000; Renick & Harter, 1989). Gleichzeitig fanden sich in Teilstudie 4 keine Hinweise dafür, dass eine separate Beschulung in Förderschulen für Schülerinnen und Schüler mit SPF mit einer Etikettierung oder Stigmatisierung einhergeht (Goffman, 1986; Tracey & Marsh, 2000). Solche sogenannten Assimilationseffekte (Elbaum, 2002; Tracey & Marsh, 2000) würden sich in Teilstudie 4 darin zeigen, dass Kinder mit SPF in Förderschulen niedrigere akademische Selbstkonzepte und weniger Lernfreude berichten als Kinder in Grundschulen, wenn für Schulartunterschiede in der mittleren Leistungsstärke der Klassen kontrolliert wird (Trautwein et al., 2006). Die wahrgenommene Leistungsstärke beziehungsweise das Prestige der besuchten Schule sollte in zukünftigen Studien jedoch direkt erfragt werden, da die Beschulungsart und mittlere Klassenleistung sehr hoch miteinander korrelieren.

ren. In einer aktuellen Studie zeigte sich, dass Kinder in Förderschulen schon in der vierten Jahrgangsstufe mehr Stigmatisierung empfinden als Kinder mit SPF in Grundschulen (Schwinger et al., 2014). Das höhere Stigmatisierungsempfinden scheint sich jedoch nicht in bedeutsamen Unterschieden im akademischen Selbstkonzept und in der Selbstwirksamkeit niederzuschlagen (Wild et al., 2015).

Während die mittlere Leistungsstärke der Klasse positiv mit der schulischen Motivation zusammenhing, trug die mittlere wahrgenommene soziale Unterstützung durch Mitschülerinnen und Mitschüler nicht zusätzlich zur Vorhersage motivationaler Unterschiede zwischen Kindern mit SPF in Förder- und Grundschulen bei. Jedoch berichteten Kinder mit SPF, die sich individuell von ihren Mitschülerinnen und Mitschülern emotional unterstützt fühlten, eine höhere schulische Motivation. Die individuell wahrgenommene Unterstützung könnte zu positiven Lerneinstellungen beitragen (Patrick et al., 2007; s. Kapitel 2.5.4.2) und eine Ressource gegenüber belastenden Situationen im Unterricht darstellen (Demaray & Malecki, 2002; Furrer & Skinner, 2003). Allerdings war die soziale Unterstützung kein bedeutsamer Prädiktor für motivationale Unterschiede zwischen Kindern mit SPF in Förder- und Grundschulen. In zukünftigen Studien könnten die sozialen Beziehungen von Kindern mit SPF detaillierter untersucht werden (Henke et al., 2017). In Teilstudie 4 wurde die globale wahrgenommene Unterstützung durch Mitschülerinnen und Mitschüler betrachtet. Jedoch könnten wechselseitige, unterstützende Beziehungen zu einzelnen Mitschülerinnen und Mitschülern für eine positive motivationale und psychosoziale Entwicklung ausreichend sein (Baumeister & Leary, 1995). Demnach sollten anstatt der sozialen Unterstützung der gesamten Klasse die sozialen Netzwerke innerhalb der Klasse in den Blick genommen werden (Wheeler & Nezlek, 1977).

5.2 Inhaltliche und methodische Herausforderungen bei der Interpretation der vorliegenden Ergebnisse

5.2.1 Auswirkungen der Datengrundlage auf die Interpretation der vorliegenden Ergebnisse

Die Aussagekraft der vorliegenden Ergebnisse wird durch die Datengrundlage mitbestimmt. Grundlage dieser Dissertation ist der IQB-Ländervergleich 2011, ein Large-Scale Assessment am Ende der vierten Jahrgangsstufe. Ziel solcher Studien ist ein vergleichendes Monitoring verschiedener Schulsysteme in Deutschland (Baumert, 2016b; Hartmann, Decristan & Klieme, 2016; Pant, 2015). Aus Large-Scale Assessments lassen sich Hinweise zu relativen Stärken und Schwächen auf Systemebene ableiten (Chudowsky & Pellegrino, 2003). Darüber hinaus können auch Zusammenhänge zwischen schulischen Outcomes und Merkmalen der Lernumwelt beziehungsweise des Schulsystems durch geeignete Auswertungsverfahren analysiert werden (Baumert, 2016b). Die Datengrundlage ist jedoch auch mit inhaltlichen Einschränkungen verbunden, die bei der Interpretation solcher Ergebnisse berücksichtigt werden müssen.

Eine erste Einschränkung besteht darin, dass die im IQB-Ländervergleich 2011 einbezogenen Schülerinnen und Schüler mit SPF nicht repräsentativ für die gesamte Schülerpopulation mit SPF in Deutschland sind. Es wurden nur Schülerinnen und Schüler einbezogen, die die Kompetenztests selbstständig bearbeiten konnten (Richter, Engelbert et al., 2012). Außerdem wurden nur Kinder berücksichtigt, die durch die Testsituation und die Testanweisungen nicht stark emotional oder geistig belastet worden wären. Ausgeschlossen wurden darüber hinaus auch Schülerinnen und Schüler, die weniger als ein Jahr Deutschunterricht hatten (Richter, Engelbert et al., 2012). Daraus lässt sich schlussfolgern, dass die vorliegende Stichprobe überwiegend aus Kindern mit SPF bestand, die leichte bis moderaten Lernbeeinträchtigungen aufweisen. Auch bildet die Datengrundlage dieser Dissertation nicht die Verteilung der Schülerpopulation mit SPF auf Förderschulen und allgemeinen Schulen ab. Die Mehrheit (ca. 66%) der Schülerstichprobe mit SPF im IQB-Ländervergleich 2011 besuchte eine allgemeine Schule. Im Gegensatz dazu wurde die Mehrheit der Schülerpopulation mit SPF 2011 in Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2016; s. a. Kapitel 2.3.3.1). Die Ergebnisse der vorliegenden Teilstudien können somit nicht auf die

gesamte Schülerpopulation mit SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* übertragen werden.

Eine zweite Einschränkung besteht darin, dass die Datengrundlage des IQB-Ländervergleichs 2011 nicht speziell auf die Schülergruppe mit SPF ausgerichtet war, sondern Aussagen über die gesamte Schülerschaft der vierten Jahrgangsstufe getroffen werden sollten. Deshalb wurden weder die Kompetenztests noch die Befragungsinstrumente speziell für die Schülergruppe mit SPF entwickelt beziehungsweise an deren Lernvoraussetzungen angepasst (s. Kapitel 2.4.1 und Kapitel 5.1.1). In Teilstudie 1 zeigte sich, dass die eingesetzten Kompetenztests geeignet waren, die schulischen Kompetenzen von Kindern mit SPF in Förderschulen und allgemeinen Schulen vergleichbar zu erfassen. In Teilstudie 4 wurde die schulische Motivation und soziale Unterstützung bei Kindern mit SPF mit den gleichen Fragebögen und unter den gleichen Bedingungen erhoben wie bei Kinder ohne SPF. Die teilweise niedrigen internen Konsistenzen der Skalen und Einschränkungen der Messinvarianz deuten darauf hin, dass die Erfassung motivationaler und psychosozialer Merkmale für die Schülergruppe mit SPF angepasst werden sollte (Byrne, 1996; Schwinger et al., 2015). So könnte auf den Einsatz negativ formulierter Items verzichtet werden und die Befragung in Einzel- beziehungsweise Kleingruppensettings durchgeführt werden (Marsh, Tracey & Craven, 2006; Nusser et al., 2015). Bei der individuellen Administration der Fragebögen könnten Testleiter besser auf Verständnisschwierigkeiten eingehen (Marsh et al., 2006).

Eine dritte Einschränkung bei der Interpretation der vorliegenden Ergebnisse liegt darin, dass keine Informationen darüber zur Verfügung stehen, wie der Unterricht für Schülerinnen und Schüler mit SPF ausgestaltet wurde und wie Schülerinnen und Schüler mit SPF in allgemeinen Schulen gefördert wurden. Neben den in den Analysen betrachteten Merkmalen der Klassenzusammensetzung könnten auch die Lehrpläne und die Unterrichtsgestaltung zu leistungsbezogenen und motivationalen Unterschieden von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen beitragen (s. Kapitel 2.5.2.3). Diese institutionellen Merkmale von Förderschulen und allgemeinen Schulen sollten in Zukunft stärker in den Blick genommen werden.

Die vorliegende Datengrundlage weist aber auch methodische Stärken auf. So stellt der IQB-Ländervergleich 2011 Daten für eine bundesweite, vergleichsweise große Stichprobe von Schülerinnen und Schülern mit SPF bereit. Außerdem erlaubt die vorliegende Datengrundlage eine nach Förderschwerpunkt differenzierte Betrachtung. Zudem wurden gesamte Schulklassen unter standardisierten Bedingungen getestet und befragt. Dadurch konnte die Klassenzu-

sammensetzung als Merkmal der Lernumgebung in Förderschulen und allgemeinen Schulen in die Analysen einbezogen werden. Die hohe Anzahl an Kindern und Schulklassen verbessert dabei die Schätzung der Zusammenhänge zwischen Merkmalen der Klassenzusammensetzung und schulischen Outcomes der Kinder mit SPF (Duncan, Jones & Moon, 1998).

5.2.2 Methodische Herausforderungen bei der Schätzung von Schularteffekten

Neben den im vorherigen Kapitel beschriebenen inhaltlichen Einschränkungen führt die querschnittliche Datengrundlage auch zu methodischen Herausforderungen bei der Schätzung von Effekten¹⁵ der Schulart und Klassenzusammensetzung (Baumert et al., 2006). Eine zentrale methodische Herausforderung in den vorliegenden Analysen sind Selektionseffekte (*Selection Bias*; Duncan, Magnuson & Ludwig, 2004). Der Besuch einer Schulart beziehungsweise bestimmten Schule erfolgt nicht zufällig, sondern Schülerinnen und Schüler werden auf Basis ihrer schulischen Leistungen beziehungsweise ihrer Leistungsfähigkeit (und im Primarbereich auf Basis des Einzugsgebietes) einer Schulart zugewiesen. Dabei wirken auch Prozesse der Selbstselektion, zum Beispiel durch die Wahl einer Schulart (oder eines bestimmten Wohngebietes). Bleibt die differenzielle Eingangsselektivität verschiedener Schularten unberücksichtigt, kann das zu einer Überschätzung institutioneller beziehungsweise kompositioneller Effekte führen (Baumert et al., 2006). In diesem Fall können potenzielle Unterschiede zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen nicht eindeutig auf die Schulart zurückgeführt werden, sondern sie könnten auch durch unterschiedliche Lernvoraussetzungen verursacht worden sein, die bereits vor der Zuweisung zu Förderschulen und allgemeinen Schulen bestanden.

Ein Beispiel für unterschiedliche Lernvoraussetzungen, die in querschnittlichen Studien häufig nicht adäquat berücksichtigt werden können, ist das bereichsspezifische Vorwissen beziehungsweise die Leistungsfähigkeit der Schülerinnen und Schüler. Diese Merkmale bilden zentrale Kriterien für die Zuweisung zu einer Schulart (Baumert et al., 2006). Zwar können die kognitiven Fähigkeiten der Schülerinnen und Schüler als Proxy für die Leistungsfähigkeit und das Vorwissen verwendet werden, allerdings wird dabei die Annahme getroffen,

¹⁵ Effekte beziehen sich dabei auf Schätzer aus statistischen Auswertungsverfahren (z. B. Regressionskoeffizienten) und nicht auf kausale Zusammenhänge zwischen Schulart und Merkmalen der individuellen schulischen Entwicklung (Rutter, 1983).

dass die kognitiven Fähigkeiten zeitlich stabil sind und nicht differenziell durch die besuchte Schulart beeinflusst werden. Diese Annahme hat sich in empirischen Untersuchungen nicht bestätigt (Becker, Lüdtke, Trautwein & Köller, 2012). Außerdem könnte die Berücksichtigung kognitiver Fähigkeiten in statistischen Analysen zu einer Unterschätzung von Schularteffekten führen. Wenn Kinder mit SPF in Grundschulen stärkere Fortschritte in ihren kognitiven Fähigkeiten erzielen als Kinder mit SPF in Förderschulen (Wocken, 2000) und die kognitiven Fähigkeiten positiv mit Merkmalen der schulischen Entwicklung zusammenhängen, dann würde man diese indirekten Effekte der Schulart kontrollieren und somit den Einfluss der Schulart auf die schulischen Leistungen unterschätzen. Diese methodische Herausforderung wird auch als Mediationsbias oder Überkontrolle (Baumert et al., 2006) bezeichnet.

Neben dem bereichsspezifischen Vorwissen und der Leistungsfähigkeit der Schülerinnen und Schüler wurden in der vorliegenden Studie weitere Hintergrundmerkmale der Schülergruppe mit SPF nicht erhoben (z. B. Schweregrad der Beeinträchtigung und Unterstützungsbedarf), die die Zuweisung zu einer Förderschule oder allgemeinen Schule beeinflussen. So werden Schülerinnen und Schüler mit einem höheren Unterstützungsbedarf eher an eine Förderschule überwiesen beziehungsweise direkt in eine Förderschule eingeschult (Sälzer et al., 2015; Wocken, 2005). Dadurch könnten Leistungsunterschiede von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen auch zum Teil auf unterschiedliche Lernvoraussetzungen und Ausgangsleistungen zurückgeführt werden, auch wenn Unterschiede in den kognitiven Fähigkeiten berücksichtigt wurden (Tiekstra, Hessels & Minnaert, 2009). In zukünftigen Studien bedarf es detaillierter Informationen zur Schülergruppe mit SPF, um unterschiedliche Lernvoraussetzungen von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen bei der Schätzung von Schularteffekten besser berücksichtigen zu können.

5.2.3 Statistische Kontrolle von Selektionseffekten durch Propensity-Score-Matching

In Schulleistungsstudien lassen sich Selektionseffekte nicht vollständig kontrollieren, da Kinder und Jugendliche nicht zufällig einer Schule beziehungsweise einer Schulart zugewiesen werden können. Zur statistischen Berücksichtigung von Selektionseffekten in Schulleistungsstudien lassen sich verschiedene Verfahren anwenden (Legewie, 2012). In den vorliegenden Analysen wurde versucht, potenzielle Unterschiede in den Lernvoraussetzungen von Kindern mit SPF in Förder- und Grundschulen durch Propensity-Score-Matching-Verfahren statistisch zu kontrollieren.

Beim Propensity-Score-Matching wird für jede Schülerin und jeden Schüler anhand der für die Zuweisung zu einer Schulart relevanten Hintergrundmerkmale eine bedingte Wahrscheinlichkeit (Propensity Score) berechnet, eine Förderschule oder eine allgemeine Schule zu besuchen. Anschließend werden nur diejenigen Schülerinnen und Schüler beider Gruppen in ihren schulischen Outcomes miteinander verglichen, die ähnliche Übergangswahrscheinlichkeiten haben. Ist das Matching erfolgreich, unterscheiden sich diese beiden Schülergruppen in Förderschulen und allgemeinen Schulen nicht mehr bedeutsam in den betrachteten Hintergrundmerkmalen (Rosenbaum & Rubin, 1985). Die Generalisierbarkeit der Gruppenvergleiche nach dem Propensity-Score-Matching beschränkt sich dabei jedoch auf Schülerinnen und Schüler mit vergleichbaren Lernvoraussetzungen und Hintergrundmerkmalen.

Durch Propensity-Score-Matching-Verfahren können Selektionseffekte bei der Auswertung besser berücksichtigt werden als in (linearen) Regressionsanalysen oder Kovarianzanalysen. Im Rahmen von Regressionsanalysen und Kovarianzanalysen kann nur eine begrenzte Anzahl von möglichen konfundierten Hintergrundmerkmalen als Kovariaten berücksichtigt werden. Die Schätzung von Schularteffekten kann bei sehr vielen Kovariaten durch Probleme der Multikollinearität und Suppression erschwert werden (Foster, 2010). Beim Propensity-Score-Matching kann hingegen eine Vielzahl an Kovariaten berücksichtigt werden, die zu einem Kennwert zusammengefasst werden. Ein weiterer Vorteil von Propensity-Score-Matching-Verfahren besteht darin, dass eine Parallelisierung der Schülergruppen in Förderschulen und allgemeinen Schulen angestrebt wird, während in Regressions- beziehungsweise Kovarianzanalysen der statistische Einfluss der Kovariaten auf die abhängige Variable (z. B. schulische Kompetenzen, schulische Motivation) kontrolliert wird.

Trotz dieser Vorteile sind Propensity-Score-Matching-Verfahren mit Herausforderungen bei der Schätzung von Schularteffekten verbunden (Foster, 2010). Eine zentrale Annahme ist,

dass die Zuweisung zu einer Schulart in der betrachteten Stichprobe nach dem Matching zufällig erfolgt (Rosenbaum & Rubin, 1983). Diese Annahme gilt nur für hinreichend große Stichprobenumfänge und für beobachtete Hintergrundmerkmale (Foster, 2010). Wie im vorherigen Kapitel beschrieben, wurden in der vorliegenden Studie bedeutsame Hintergrundmerkmale (z. B. bereichsspezifisches Vorwissen) nicht explizit berücksichtigt. Dadurch kann die Schätzung von Schularteffekten auch nach erfolgreichem Matching durch unbeobachtete Faktoren verzerrt werden (Schafer & Kang, 2008). Zum einen könnte der Einfluss der Schulart überschätzt werden, da die Schülergruppen in querschnittlichen Studien nicht hinsichtlich ihres Ausgangsniveaus parallelisiert werden können. Zum anderen könnten Schularteffekte aber auch unterschätzt werden, wenn die berücksichtigten Hintergrundmerkmale durch die Schulart selbst beeinflusst werden (z. B. kognitive Fähigkeiten, Becker et al., 2012); in diesem Fall wird ein Teil des Schularteffekts herausgerechnet (Baumert et al., 2006). Zusammenfassend kann festgehalten werden, dass durch Propensity-Score-Matching-Verfahren für querschnittliche Datengrundlagen relativ genaue Schätzungen von Schularteffekten möglich sind. Allerdings sind diese Schätzungen im Vergleich zu längsschnittlichen Studien mit einer höheren Unsicherheit behaftet, da unterschiedliche Ausgangsleistungen nicht berücksichtigt werden können.

5.3 Diskussion der Zusammenhänge zwischen Beschulungsart und schulischen Outcomes von Kindern mit sonderpädagogischem Förderbedarf

Die zentrale Fragestellung der Dissertation war, wie die Beschulung in Förderschulen und allgemeinen Schulen mit den schulischen Outcomes von Schülerinnen und Schülern mit SPF zusammenhängt. Die Beantwortung dieser Frage kann Hinweise darauf geben, in welcher Schulart Schülerinnen und Schüler mit SPF in ihrer schulischen Entwicklung stärker profitieren. Allerdings weisen die Ergebnisse der zweiten, dritten und vierten Teilstudie auf gegenläufige Zusammenhänge zwischen Beschulungsart und schulischen Kompetenzen beziehungsweise schulischer Motivation von Schülerinnen und Schülern mit SPF hin. Auf der einen Seite hatten Kinder mit SPF in Grundschulen höhere schulische Kompetenzen als vergleichbare Kinder mit SPF in Förderschulen (Teilstudien 2 und 3). Auf der anderen Seite berichteten Kinder mit SPF in Förderschulen eine höhere schulische Motivation (akademisches Selbstkonzept, Lernfreude) als Kinder mit SPF in Grundschulen (Teilstudie 4).

Diese Ergebnisse stehen auf den ersten Blick im Widerspruch zu querschnittlichen und längsschnittlichen Studien, die positive Zusammenhänge zwischen individuellen schulischen Leistungen und akademischem Selbstkonzept beziehungsweise Lernfreude berichten (Grolnick & Ryan, 1990; Köller, Trautwein, Lüdtke & Baumert, 2006; Marsh et al., 2008; Valentine et al., 2004). Aufgrund der Leistungsvorteile zugunsten von Kindern und Jugendlichen mit SPF in allgemeinen Schulen könnte man erwarten, dass diese Schülergruppe ein höheres akademisches Selbstkonzept beziehungsweise mehr Lernfreude als Schülerinnen und Schüler mit SPF in Förderschulen berichten. Allerdings lassen sich auch für die Schülergruppe ohne SPF vergleichbare Schularteffekte auf schulische Leistungen und Merkmale der schulischen Motivation feststellen (Baumert et al., 2006; Köller, 2004; Stäbler, Dumont, Becker & Baumert, 2017; Wouters, De Fraine, Colpin, Van Damme & Verschueren, 2012). Die gegenläufigen Zusammenhänge zwischen Beschulungsart und schulischen Leistungen beziehungsweise schulischer Motivation zeigen sich auch in den differenziellen Effekten des mittleren Leistungsniveaus einer Klasse (Baumert et al., 2006; Hanushek, Kain, Markman & Rivkin, 2003; Justice et al., 2014; Marsh et al., 2008). Während das mittlere Leistungsniveau positiv mit der individuellen Leistungsentwicklung zusammenhängt, zeigen sich negative Effekte auf das akademische Selbstkonzept und weitere motivationale Merkmale (Köller, 2004; Stäbler et al., 2017).

Auf der einen Seite kann sich die mittlere Klassenleistung positiv auf die individuelle Leistungsentwicklung auswirken (s. Kapitel 2.5.2.2). So stellen leistungstärkere Schularten beziehungsweise Klassen eine anregende Lernumgebung zur Verfügung, die durch hohe Erwartungshaltung der Lehrkräfte und Interaktionen mit leistungsstarken Mitschülerinnen und Mitschülern gekennzeichnet ist (Dreeben & Barr, 1988; Haeberlin, 1991; Hornstra et al., 2010; Köller et al., 2013). Dabei können leistungsstarke Mitschülerinnen und Mitschüler als Lernvorbilder für Schülerinnen und Schüler mit SPF fungieren (Bandura, 1971; Farrell, Dyson, Polat, Hutcheson & Gallannaugh, 2007; Justice et al., 2014; Wilkinson, 2002). Außerdem könnten die Lehrkräfte in leistungstärkeren Klassen einen anspruchsvolleren Unterricht durchführen, von dem Schülerinnen und Schüler mit SPF in ihrer Leistungsentwicklung profitieren (Wocken, 2005).

Auf der anderen Seite wirkt sich die Leistungsstärke der Schulart beziehungsweise Klasse negativ auf das akademische Selbstkonzept und die Lernfreude von Schülerinnen und Schülern mit SPF aus (s. Kapitel 2.5.4). Insbesondere das akademische Selbstkonzept wird stärker durch soziale Vergleiche mit Mitschülerinnen und Mitschülern in der Klasse beeinflusst

als durch Vergleiche mit Kindern und Jugendlichen anderer Schularten (Marsh, 2005; Trautwein et al., 2006). So fand zum Beispiel Köller (2004), dass Schülerinnen und Schüler in leistungsstärkeren Schularten beziehungsweise Klassen niedrigere Fähigkeitsselbsteinschätzungen und ein geringeres fachliches Interesse berichten. Trotz dieser motivationalen Unterschiede erreichten Schülerinnen und Schüler in Gymnasien auch nach statistischer Kontrolle der Lernvoraussetzungen und des Vorwissens höhere schulische Kompetenzen als Schülerinnen und Schüler an anderen Schularten (Köller, 2004).

Eine längsschnittliche Betrachtung gegenläufiger Effekte der mittleren Leistungsstärke innerhalb eines Schuljahres wurde von Stäbler et al. (2017) vorgenommen. Sie fanden, dass die mittlere Klassenleistung einen positiven Einfluss auf die individuelle Leistung von Jugendlichen ohne SPF zur Mitte und zum Ende des Schuljahres hat (Stäbler et al., 2017). Gleichzeitig zeigte sich nur zu Beginn des Schuljahres ein negativer Einfluss der mittleren Klassenleistung auf das akademische Selbstkonzept. Im Verlauf des Schuljahres konnte kein zusätzlicher Effekt der mittleren Klassenleistung auf das akademische Selbstkonzept beobachtet werden (Stäbler et al., 2017). Somit ließen sich Fischteicheffekte nur im Querschnitt, jedoch nicht im Längsschnitt nachweisen (s. a. Marsh et al., 2008). Eine mögliche Erklärung dafür könnte sein, dass Kinder und Jugendliche sich zu Beginn des Schuljahres mit ihren Mitschülerinnen und Mitschülern vergleichen und auf dieser Grundlage ihr akademisches Selbstkonzept ausbilden. Diese Fähigkeitseinschätzungen bleiben im Verlauf des Schuljahres dann konstant (Stäbler et al., 2017).

Auch die Studie von Wouters et al. (2012) weist darauf hin, dass sich die Leistungsstärke der Klasse unmittelbar nach dem Wechsel in eine neue Schulart auf das akademische Selbstkonzept auswirkt. Schülerinnen und Schüler, die von einem leistungsstärkeren Bildungsgang in einen leistungsschwächeren Bildungsgang wechselten, verzeichneten einen anfänglichen Anstieg im akademischen Selbstkonzept (im Vergleich zu Schülerinnen und Schülern, die im selben Bildungsgang verblieben). Die Ergebnisse von Stäbler et al. (2017) und Wouters et al. (2012) könnten in zukünftigen Studien innerhalb eines Studiendesigns untersucht werden. Während bei Stäbler et al. (2017) keine Schularteffekte betrachtet wurden, stand Wouters et al. (2012) kein direktes Maß für die mittlere Leistungsstärke der Klasse zur Verfügung. Eine gemeinsame Betrachtung der Schulart und der leistungsbezogenen Schülerzusammensetzung ermöglicht die Analyse von Interaktionen zwischen institutionellen und kompositionellen Merkmalen der Lernumgebung (s. Kapitel 2.5.2.4). Dadurch könnten die gegenläufigen

Zusammenhänge der Beschulungsart mit den schulischen Kompetenzen und der schulischen Motivation von Schülerinnen und Schülern mit SPF detaillierter untersucht werden.

Des Weiteren könnte die Richtung sozialer Vergleiche die gegenläufigen Zusammenhänge zwischen Beschulungsart und schulischen Leistungen beziehungsweise akademischem Selbstkonzept erklären (Blanton, Buunk, Gibbons & Kuyper, 1999; Dijkstra et al., 2008; Köller, 2004). Dabei lassen sich soziale Aufwärtsvergleiche mit leistungstärkeren Mitschülerinnen und Mitschülern und soziale Abwärtsvergleiche mit leistungsschwächeren Schülerinnen und Schülern unterscheiden (Corcoran et al., 2011). Welche Richtung sozialer Vergleiche Schülerinnen und Schüler präferieren, hängt von ihren Motiven und auch von der Leistungsstärke der Lerngruppe ab (Blanton et al., 1999; Dijkstra et al., 2008). Die Richtung sozialer Vergleiche hat zudem unterschiedliche Auswirkungen auf die schulischen Leistungen und das akademische Selbstkonzept (Collins, 1996; Dijkstra et al., 2008). Abwärtsvergleiche mit Leistungsschwächeren werden unter anderem aus Motiven der Selbstwerterhöhung angestrengt und führen zu einer Steigerung des akademischen Selbstkonzepts (Collins, 1996). Jedoch werden soziale Abwärtsvergleiche seltener aufgrund von Motiven der Selbstverbesserung durchgeführt und könnten dadurch zu geringeren Leistungsverbesserungen führen (Altermatt & Pomerantz, 2005; Dai & Rinn, 2008).

Im Gegensatz dazu werden soziale Aufwärtsvergleiche von Motiven der Selbstverbesserung begleitet und können dadurch eine stärkere Leistungsentwicklung anregen (Blanton et al., 1999). Allerdings gehen Aufwärtsvergleiche mit niedrigeren akademischen Selbstkonzepten einher, da eigene Leistungsdefizite stärker sichtbar werden (Blanton et al., 1999; Dijkstra et al., 2008). In leistungstärken Lerngruppen treten soziale Aufwärtsvergleiche häufiger auf als in leistungsschwächeren Lerngruppen (Köller, 2004). So könnten allgemeine Schulen im Vergleich zu Förderschulen bei Kindern und Jugendlichen mit SPF stärker soziale Aufwärtsvergleiche anregen und dadurch zu leistungsbezogenen Unterschieden zugunsten von allgemeinen Schulen und motivationalen Unterschieden zugunsten von Förderschulen beitragen (Tracey & Marsh, 2000).

Zusammenfassend scheint der Besuch einer allgemeinen Schule mit Leistungsvorteilen für Schülerinnen und Schüler mit SPF einherzugehen. Trotz der geringeren schulischen Motivation erzielen Kinder mit SPF in allgemeinen Schulen deutlich höhere schulische Kompetenzen als Kinder mit SPF in Förderschulen. Insbesondere könnten Kinder und Jugendliche mit SPF von leistungstärkeren Lerngruppen in allgemeinen Schulen profitieren, auch wenn sich die mittlere Leistung der Lerngruppe in Form von Referenzgruppeneffekten negativ auf akademi-

sches Selbstkonzept und Lernfreude auswirken kann. Diese gegenläufigen Zusammenhänge der mittleren Leistungsstärke einer Klasse zeigen sich auch bei Schulartvergleichen mit Schülerinnen und Schülern ohne SPF (Baumert et al., 2006; Dijkstra et al., 2008; Köller, 2004; Wouters et al., 2012). Es werden jedoch längsschnittliche Studien benötigt, um den Einfluss der Lernumgebung auf die kognitive und motivationale Entwicklung bei Schülerinnen und Schülern mit SPF und deren Zusammenspiel zu untersuchen (Bakker et al., 2007; Gorges, Neumann, Wild, Stranghöner & Lütje-Klose, 2017). In einer aktuellen Studie wurden die längsschnittlichen Zusammenhänge zwischen individueller Leistung und akademischem Selbstkonzept von Kindern mit dem Förderschwerpunkt *Lernen* in Förder- und Grundschulen in den Blick genommen (Gorges et al., 2017). Dabei zeigten sich im Zeitverlauf von der dritten zur vierten Klasse nicht zu allen drei Messzeitpunkten reziproke Effekte zwischen schulischer Leistung und akademischem Selbstkonzept (Gorges et al., 2017). Bedeutsame Effekte ließen sich auch nur von schulischen Leistungen auf das akademische Selbstkonzept feststellen. Diese Ergebnisse deuten darauf hin, dass sich etablierte theoretische Annahmen zum Zusammenhang zwischen Leistung und Motivation nicht ohne weiteres auf die Schülergruppe mit SPF übertragen lassen. So könnten unterschiedliche Feedbackstrukturen und Lernziele auf intraindividuelle Ebene zu einer Entkopplung von schulischen Leistungen und akademischem Selbstkonzept bei Kindern mit SPF führen (Grolnick & Ryan, 1990; Nusser & Wolter, 2016).

Bei der Evaluation der Lernumgebung sollten sowohl kognitive als auch motivationale und psychosoziale Merkmale herangezogen werden. Hier stellt sich die Frage, welche schulischen Outcomes bei Platzierungsentscheidungen eine wichtigere Rolle beigemessen werden sollte. Auch wenn in den vorliegenden Teilstudien gegenläufige Zusammenhänge der Beschulungsart mit den schulischen Leistungen und der schulischen Motivation von Kindern mit SPF festgestellt wurden, kann die Förderung beider schulischer Outcomes prinzipiell vereinbar sein (Deci & Chandler, 1986; Helmke & Weinert, 1997; Grosche, 2015). So könnte die schulische Motivation von Schülerinnen und Schülern mit SPF in allgemeinen Schulen durch individuelle Bezugsnormen bei der Leistungsrückmeldung und ein unterstützendes Klassenklima gefördert werden (Allodi, 2000; Chapman, 1988; Deci & Chandler, 1986; Forman, 1988).

Platzierungsentscheidungen sollten auch unter Berücksichtigung zukünftiger Bildungschancen der Schülergruppe mit SPF getroffen werden. Trotz der niedrigeren schulischen Motivation weisen Schülerinnen und Schüler mit SPF in allgemeinen Schulen nicht nur höhere

schulische Kompetenzen auf, sie erhalten auch höheren Bildungschancen (Pfahl, 2012). Die Mehrheit der Förderschülerinnen und -schüler beendet ihre Schulkarriere ohne einen Schulabschluss, der eine berufliche Ausbildung oder weiterführende Bildung ermöglichen würde (Powell & Pfahl, 2012). In manchen Bundesländern kann in Förderschulen nicht einmal ein Hauptschulabschluss erreicht werden (Pfahl, 2012). Dies erschwert den Einstieg ins Berufsleben (Pfahl, 2012). Auch wenn zu den berufsqualifizierenden Abschlüssen der Schülergruppe mit SPF in allgemeinen Schulen noch keine verlässlichen Aussagen getroffen werden können (Malecki, 2013; KMK, 2016), scheinen die Bildungschancen dieser Schülergruppe höher zu sein. Im Gegensatz dazu besteht im Zuge der Bildungsexpansion (Bos et al., 2010; Solga, 2005) für Schülerinnen und Schüler mit einem Förderschulabschluss die Gefahr, vom Arbeitsmarkt zunehmend ausgeschlossen zu werden (Powell & Pfahl, 2012). So wird in der Salamanca-Erklärung (1994) darauf hingewiesen, dass eine gemeinsame Beschulung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen zu einem Lernniveau beitragen kann, das eine Teilhabe am beruflichen und gesellschaftlichen Leben ermöglicht. Insbesondere für Kinder und Jugendliche mit den Förderschwerpunkten *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* werden Fragen der Platzierung in Förderschulen oder allgemeinen Schulen deshalb auch im Hinblick auf Aspekte der Bildungsgerechtigkeit diskutiert (Moser, 2012; Preuss-Lausitz, 2013).

5.4 Implikationen für zukünftige Forschung

In diesem Kapitel werden Empfehlungen zusammengefasst, welche Richtungen empirische Forschung zum Themenfeld Inklusion und Schülerinnen und Schüler mit SPF einschlagen könnte. Im ersten Teil wird dafür plädiert, dass zukünftige Studien die zugrundeliegenden Prozesse und Erklärungsansätze für mögliche Schulartunterschiede untersuchen sollten. Im zweiten Teil werden zwei Forschungsprojekte vorgestellt, die den Fokus der vorliegenden Arbeit erweitern. Erstens soll in weiterführenden Analysen von Daten des IQB-Ländervergleichs 2011 geprüft werden, wie der gemeinsame Unterricht von Kindern mit SPF in allgemeinen Schulen mit den schulischen Leistungen, Merkmalen der schulischen Motivation und der sozialen Integration von Kindern *ohne* SPF zusammenhängt. Zweitens wird in einer längsschnittlich angelegten Studie untersucht, inwieweit der festgestellte Förderschwerpunkt zeitlichen Veränderungen unterliegt und wie diese Veränderungen mit der Entwicklung im Leseverständnis zusammenhängt.

5.4.1 Forschungsbedarfe für zukünftige Studien

In den Teilstudien dieser Dissertation wurden Unterschiede zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen in den schulischen Kompetenzen und Merkmalen der schulischen Motivation untersucht. Außerdem wurden Zusammenhänge zwischen diesen schulischen Outcomes mit der Klassenkomposition als mögliche Erklärungen für Schulartunterschiede betrachtet. Zukünftige Studien sollten verstärkt die zugrundeliegenden Prozesse für diese Schulartunterschiede in den Blick nehmen (Deutsche Gesellschaft für Erziehungswissenschaft, DGfE, 2017). Auch besteht Forschungsbedarf, wie sich die Inklusion von Schülerinnen und Schülern mit SPF in allgemeine Schulen so umsetzen lässt, dass alle Schülerinnen und Schüler in ihrer schulischen Entwicklung profitieren (Dyson, 2014). Dazu werden längsschnittliche und experimentelle Studien benötigt, um den Einfluss von Merkmalen der Lernumgebung auf individuelle schulische Outcomes quantifizieren und zugrundeliegende Prozesse identifizieren zu können (Baumert, 2016b; Klieme & Vieluf, 2013; Pant, 2015; Tillmann, 2016). Dabei sollte die konkrete Ausgestaltung von Inklusion im Unterricht beleuchtet werden. Zur konkreten Ausgestaltung von Inklusion im Unterricht gehören zum Beispiel Unterrichtsmerkmale, die die Heterogenität der gesamten Schülerschaft angemessen berücksichtigen (Prengel, 2013). So werden Maßnahmen der Binnendifferenzierung und eine auf individuellen Lernzielen basierende Bewertung als wichtige Aspekte inklusiver Unterrichtsgestaltung angesehen, die in zukünftigen Studien untersucht werden sollten (Grosche, 2015; Hinz, 2002).

In dieser Dissertation wurden Unterschiede in den schulischen Kompetenzen und Merkmalen der schulischen Motivation zwischen Kindern mit SPF in Förder- und Grundschulen betrachtet. Neben der effektiven Leistungsförderung aller Schülerinnen und Schüler besteht ein wichtiges Ziel von Inklusion darin, die Anerkennung und soziale Teilhabe von Schülerinnen und Schülern mit SPF zu fördern (Huber, 2009; Prengel, 2013). In zukünftigen Studien sollten deshalb auch Merkmale der sozialen Integration und des schulischen Wohlbefindens von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen berücksichtigt werden. Die Förderung der psychosozialen Entwicklung ist dabei nicht nur ein Ziel inklusiver Bildung an sich, sondern kann auch zu schulischen Erfolgen von Kindern und Jugendlichen mit SPF beitragen (Deci & Chandler, 1986).

Weiterhin besteht Forschungsbedarf zu Fragen der Zuweisung von Kindern und Jugendlichen mit SPF zu Förderschulen oder allgemeinen Schulen. In Kapitel 2.3.2 wurde dargestellt, dass

bei der Feststellung eines SPF und der anschließenden Empfehlung für eine Förderschule beziehungsweise allgemeine Schule unterschiedliche Verfahren zum Einsatz kommen, die sich zwischen den Ländern unterscheiden (Kottmann, 2006; Sälzer et al., 2015). Eine umfassende und nach Förderschwerpunkt differenzierte Betrachtung, welche Merkmale auf Schüler- und Schulebene mit der Zuweisung zu einer Schulart zusammenhängen, liegt bisher noch nicht vor. Vorangegangene Studien konnten zeigen, dass Schülerinnen und Schüler in Förderschulen ungünstigere Lernvoraussetzungen haben und über geringere soziokulturelle Ressourcen in ihren Familien verfügen (Autorengruppe Bildungsberichterstattung, 2012; Klein, 2001; Wocken 2005). Allerdings wurden in der Mehrzahl dieser Studien Schülerinnen und Schüler mit SPF in Förderschulen mit der gesamten Schülergruppe in allgemeinen Schulen verglichen. Eine differenzierte Analyse der Schülerinnen und Schüler mit SPF in Förderschulen und allgemeinen Schulen könnte helfen, Selektionseffekte in zukünftigen Studien besser statistisch zu kontrollieren.

5.4.2 Forschungsausblick auf zwei Anschlussprojekte

5.4.2.1 Auswirkungen des gemeinsamen Unterrichts auf Kinder ohne sonderpädagogischen Förderbedarf

In einer ersten Anschlussstudie wird der Frage nachgegangen, ob Schülerinnen und Schüler mit SPF von einer gemeinsamen Beschulung in allgemeinen Schulen anstatt in Förderschulen profitieren können. Es stellt sich jedoch auch die Frage, wie sich eine gemeinsame Beschulung auf die schulische Entwicklung von Schülerinnen und Schülern *ohne* SPF auswirkt. Deshalb wird in einem Anschlussprojekt analysiert, wie der gemeinsame Unterricht von Kindern mit SPF in allgemeinen Schulen mit den schulischen Leistungen, Merkmalen der schulischen Motivation und der sozialen Integration von Kindern ohne SPF zusammenhängt (Kocaj, Kuhl & Stanat, 2016).

Die Auswirkungen des gemeinsamen Unterrichts auf Kinder und Jugendliche mit SPF werden aus pädagogischer Sicht kontrovers diskutiert. Auf der einen Seite wird befürchtet, dass Schülerinnen und Schüler ohne SPF durch die Anwesenheit von Schülerinnen und Schülern mit SPF in der Klasse in ihrer schulischen Entwicklung gebremst werden könnten. So könnten Schülerinnen und Schüler mit SPF im gemeinsamen Unterricht mehr Aufmerksamkeit der Lehrkräfte in Anspruch nehmen und ihre Mitschülerinnen und Mitschüler ohne SPF daher

weniger Unterstützung erhalten (Gottfried, 2014). Außerdem könnten die Lehrkräfte das Anspruchsniveau und Unterrichtstempo an Schülerinnen und Schülern mit SPF anpassen, was zu niedrigeren schulischen Kompetenzen und einer geringeren Motivation bei Kindern und Jugendlichen ohne SPF führen könnte (Ruijs, van der Veen & Peetsma, 2010). Auf der anderen Seite wird argumentiert, dass in inklusiven Klassen verstärkt binnendifferenzierende Unterrichtsmethoden eingesetzt werden, von denen auch Schülerinnen und Schüler ohne SPF profitieren sollten (Ruijs et al., 2010). Und schließlich könnten Schülerinnen und Schüler ohne SPF durch kooperative Lernformen eigene Wissensstände festigen und dadurch höhere schulische Leistungen erzielen (Fuchs, Fuchs, Mathes & Simmons, 1997).

Vor dem Hintergrund dieser Diskussionen wird in diesem Anschlussprojekt auf Grundlage der Daten des IQB-Ländervergleichs 2011 (s. Kapitel 3.2) der Frage nachgegangen, wie der gemeinsame Unterricht mit Merkmalen der schulischen Entwicklung von Schülerinnen und Schülern ohne SPF zusammenhängt. Dabei soll nach Möglichkeit auch geprüft werden, ob sich differenzielle Befundmuster für die Förderschwerpunkte *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* zeigen. In den Analysen wurden Unterschiede zwischen Integrationsklassen (Anwesenheit von mindestens einem Kind mit SPF) und Regelklassen (keine Kinder mit SPF in der Klasse) in den schulischen Kompetenzen, Merkmalen der schulischen Motivation und in der sozialen Integration untersucht, wobei Kovariaten auf Individual- und Klassenebene berücksichtigt wurden.

Vorläufige Ergebnisse sprechen dafür, dass der gemeinsame Unterricht insgesamt nicht mit negativen Auswirkungen für Kinder ohne SPF verbunden ist. Allerdings fanden sich Hinweise darauf, dass die Inklusion von Schülerinnen und Schülern mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* mit Herausforderungen für Kinder ohne SPF verbunden sein könnte. Hier wiesen Kinder ohne SPF in Integrationsklassen niedrigere Mathematikleistungen auf als Kinder ohne SPF in Regelklassen und berichteten zudem weniger Lernfreude und mehr Langeweile. Die negativen Effekte für diesen Förderschwerpunkt waren jedoch gering. Möglicherweise führen Verhaltensauffälligkeiten der Kinder mit dem Förderschwerpunkt *emotionale und soziale Entwicklung* zu Störungen im Unterricht, die das Unterrichtserleben der Mitschülerinnen und Mitschüler beeinflussen (Gottfried, 2014). In weiterführenden Analysen soll geprüft werden, ob sich die Effekte der gemeinsamen Beschulung für leistungsschwächere und leistungstärkere Kinder ohne SPF unterscheiden.

5.4.2.2 Zeitliche Veränderungen im Förderschwerpunkt und ihre Zusammenhänge mit der Leistungsentwicklung im Leseverständnis bei Grundschulkindern

In dieser Dissertation wird der individuell festgestellte Förderschwerpunkt als zeitlich stabiles Konstrukt betrachtet und aufgrund der querschnittlichen Datengrundlage nur einmal erhoben. Diese Sichtweise wird auch in längsschnittlichen Studien eingenommen. Dabei gibt es Hinweise darauf, dass die Feststellung eines SPF zeitlichen Veränderungen unterliegt: Zum einen können Schülerinnen und Schüler mit SPF im Verlauf ihrer Schulkarriere ihren Förderschwerpunkt wechseln (Kottmann, 2006; Morgan et al., 2011); zum zweiten können Schülerinnen und Schüler ihren Förderstatus nach erfolgreicher Förderung verlieren (Morgan et al., 2011); zum dritten unterscheiden sich Schülerinnen und Schüler darin, wann erstmals ein SPF bei ihnen festgestellt wurde (Sälzer et al., 2015).

In einem zweiten Anschlussprojekt wird zum einen analysiert, ob ein individuell festgestellter Förderbedarf zeitlichen Veränderungen unterliegt und wie häufig Wechsel des Förderschwerpunktes auftreten. Zum anderen wird untersucht, ob zeitliche Veränderungen im Förderschwerpunkt mit unterschiedlichen Leistungsentwicklungen im Leseverständnis der Schülerinnen und Schüler mit SPF zusammenhängen. Auf der einen Seite könnte eine frühe Feststellung mit einer erhöhten Leistungsentwicklung einhergehen, da frühzeitig Förderprogramme und Ressourcen bereitgestellt werden (Morgan et al., 2011). Im Gegensatz dazu würde eine spätere Feststellung zu einer geringeren Leistungsentwicklung führen, da kompensatorische Fördermaßnahmen mit zeitlicher Verzögerung zum Einsatz kommen (Wei et al., 2011). Auf der anderen Seite wird argumentiert, dass eine spätere Feststellung Etikettierungseffekten entgegenwirkt und dadurch einen positiven Einfluss auf die Leistungsentwicklung der Schülerinnen und Schüler hat (Pfahl, 2012).

Grundlage der Studie sind Daten des *Reading First* Programms in Michigan (USA) von 2002 bis 2008 (Carlisle, Cortina & Zeng, 2010). Dabei handelt es sich um eine staatlich geförderte Intervention zur Förderung des Leseverständnisses von sozial benachteiligten Kindern in den ersten Schuljahren. Die Analysestichprobe bestand aus $N = 10\,196$ Kindern aus $n = 191$ Schulen und drei Kohorten, die zu drei Messzeitpunkten (1. Klasse bis 3. Klasse) im jährlichen Abstand getestet wurden. Bei $n = 1193$ Kindern wurde im Verlauf der ersten drei Schuljahre eine spezifische Lernbehinderung beziehungsweise eine sprachliche Beeinträchtigung festgestellt. Für jedes der drei Schuljahre lagen Informationen zum Förderschwerpunkt der Schülerinnen und Schüler vor. Auf dieser Grundlage bildeten wir Schülergruppen, für die

ein unterschiedlicher zeitlicher Verlauf des SPF dokumentiert war (z. B. späte vs. frühe Feststellung eines SPF, Wechsel im Förderschwerpunkt). Erste Analysen zeigen, dass nur für einen geringen Anteil der Kinder mit SPF eine persistente Lernbehinderung beziehungsweise persistente sprachliche Beeinträchtigung festgestellt wurde. Die Mehrheit der Kinder erhielt eine späte Diagnose. Wechsel des Förderschwerpunktes kamen relativ selten vor und konnten nur von einer sprachlichen Beeinträchtigung hin zu einer Lernbehinderung festgestellt werden.

Die Leistungsentwicklung der Kinder analysierten wir anhand von linearen Wachstumsmodellen in einem Mehrebenenansatz, wobei wir zeitinvariante Kovariaten auf Schülerebene berücksichtigten. Im Hinblick auf das Ausgangsniveau und die Leistungsentwicklung ergaben sich sowohl zwischen den Schülergruppen mit sprachlichen Beeinträchtigungen als auch zwischen den Schülergruppen mit einer spezifischen Lernbehinderung keine signifikanten Unterschiede. Allerdings hatte die Schülergruppe, die von einer Sprachbeeinträchtigung zu einer spezifischen Lernbehinderung wechselte, ein signifikant geringeres Ausgangsniveau und signifikant geringere Leistungszuwächse im Leseverständnis im Vergleich zu den Schülergruppen mit einer sprachlichen Beeinträchtigung. Darüber hinaus wiesen Schülergruppen mit sprachlichen Beeinträchtigungen höhere Ausgangsleistungen auf und erzielten höhere Leistungsgewinne als Schülergruppen mit einer Lernbeeinträchtigung. Zusammenfassend weisen die vorläufigen Ergebnisse darauf hin, dass es sich bei den Förderschwerpunkten einer Lernbehinderung und einer sprachlichen Beeinträchtigung nicht zwangsläufig um zeitlich stabile Merkmale handelt und mit Veränderungen zu rechnen ist. Deshalb sollte der Förderschwerpunkt in längsschnittlichen Studien mehrmals erhoben werden. Allerdings zeigte sich lediglich für die Kinder, die ihren Förderschwerpunkten wechselten, signifikante Unterschiede in Ausgangsniveau und Leistungsentwicklung. Diese Schülergruppe wies die niedrigste Leistungsentwicklung aller betrachteten Gruppen auf, was eventuell auf die mangelnde Wirksamkeit der ursprünglichen Förderung zurückführbar ist.

5.5 Implikationen für die Bildungspolitik und pädagogische Praxis

Schulleistungsstudien wurden in den letzten 20 Jahren systematisch ausgebaut, um auch praktische Implikationen aus deren Studienergebnissen ableiten zu können (Baumert, 2016b; Tillmann, 2016). Schulleistungsstudien können zum einen Impulse für bildungspolitische und schulstrukturelle Entscheidungen, wie beispielsweise den Abbau eines separaten Förderschulsystems, liefern, zum anderen könnten Implikationen für die pädagogische Praxis abgeleitet werden (Tillmann, 2016). Dabei muss beachtet werden, dass Akteure der Bildungsforschung, Bildungspolitik und Bildungspraxis unterschiedliche Prozesse der Erkenntnisgewinnung anwenden, die auch durch institutionell verankerte Motive und Zielstellungen geprägt werden (Luhmann & Schorr, 1988). Diese Unterschiede beeinflussen die Übertragung wissenschaftlicher Ergebnisse in Handlungsanleitungen für die Bildungspolitik und pädagogische Praxis einerseits, aber auch die Übertragung bildungspraktischer und bildungspolitischer Herausforderungen in wissenschaftliche Fragestellungen andererseits (Luhmann & Schorr, 1988). Deshalb kann nicht erwartet werden, dass die Ergebnisse dieser Dissertation (und weiterer empirischer Studien) direkt in politische Entscheidungen zur Gestaltung von Bildungssystemen einfließen. Im besten Fall können die Ergebnisse aus Schulleistungsstudien zur Unterstützung politischer und bildungspraktischer Entscheidungen herangezogen werden (Baumert, 2016b).

5.5.1 Implikationen für die Bildungspolitik

Die Ergebnisse der vorliegenden Teilstudien weisen darauf hin, dass Kinder mit einem SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* in ihren schulischen Leistungen vom Besuch einer allgemeinen Schule profitieren. In Einklang mit den Ergebnissen werden Schülerinnen und Schüler mit diesen Förderschwerpunkten zunehmend in allgemeinen Schulen anstatt in Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2016; s. Kapitel 2.3.3.1). Allerdings berichteten Schülerinnen und Schüler mit SPF in Förderschulen eine höhere schulische Motivation als vergleichbare Schülerinnen und Schüler mit SPF in allgemeinen Schulen. Diese gegenläufigen Zusammenhänge der Beschulungsart mit den schulischen Leistungen und der schulischen Motivation von Schülerinnen und Schüler mit SPF erschweren die Ableitung von Implikationen für die Bildungspolitik. Welche Schlussfolgerungen aus den vorliegenden Ergebnissen hinsichtlich

des Förderortes von Schülerinnen und Schülern mit SPF getroffen werden, hängt davon ab, welchen Stellenwert der Förderung schulischer Kompetenzen und der Förderung schulischer Motivation beigemessen wird.

Die Debatte um Inklusion im Bildungsbereich wird mit empirischen und normativen Argumenten geführt (Emmerich, 2016; Moser, 2012; Veber, 2016). Die Ergebnisse dieser Dissertation bilden gemeinsam mit anderen Studien und Forschungssynthesen (z. B. Haeberlin, 1991; Lehmann & Hoffmann, 2009; Preuss-Lausitz, 2014; Spörer, Schröder-Lenzen, Vock & Maaz, 2015; Wild et al., 2015; Wocken, 2005) empirische Argumentationsgrundlagen. Die empirische Grundlage weist darauf hin, dass allgemeine Schulen im Vergleich zu Förderschulen die schulischen Kompetenzen der Kinder und Jugendlichen mit SPF effektiver fördern und der Besuch einer Förderschule soziale Disparitäten verstärken kann (Klein, 2001; Wocken, 2005). So erzielen Schülerinnen und Schüler mit SPF in allgemeinen Schulen in der Mehrzahl der Studien höhere schulische Leistungen und höhere Schulabschlüsse als Schülerinnen und Schüler mit SPF in Förderschulen (Powell & Pfahl, 2012).

Darüber hinaus sind Schülerinnen und Schüler aus Familien mit niedrigem sozialen und kulturellen Kapital in Förderschulen (insbesondere mit dem Förderschwerpunkt *Lernen*) überrepräsentiert (Autorengruppe Bildungsberichterstattung 2010; Klein, 2001; Kornmann, 2006; Wocken, 2005). Auch in den vorliegenden Analysen in Teilstudie 2 und 3 zeigte sich, dass Schülerinnen und Schüler mit SPF in allgemeinen Schulen sowohl über günstigere familiäre Bedingungen und mehr Ressourcen verfügten als Kinder in Förderschulen. Zusätzlich ist die Lernumgebung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen durch eine leistungsanregende Klassenzusammensetzung gekennzeichnet. Dadurch können sich herkunftsbedingte Bildungsnachteile von Kindern und Jugendlichen in Förderschulen im Zeitverlauf verstärken (Bos et al., 2010; Schümer, 2004). Diese Ergebnisse amtlicher Schulstatistiken und empirischer Studien sprechen dafür, Schülerinnen und Schüler mit SPF vermehrt in allgemeinen Schulen zu unterrichten. Es ist aber eine offene Frage, ob inklusive Bemühungen zu mehr Bildungsgerechtigkeit führen. Ungleiche Bildungschancen könnten auch dann auftreten, wenn alle Schülerinnen und Schüler mit SPF allgemeine Schulen besuchen. So dürften Schülerinnen und Schüler mit SPF in einem differenzierenden Schulsystem auch ohne Förderschulen eher in niedrigen Bildungsgängen unterrichtet werden (Emmerich, 2016). Es müsste also empirisch gezeigt werden, dass eine separate Beschulung von Schülerinnen und Schülern mit SPF in Förderschulen zu unterschiedlichen Bildungs- und

Teilhabechancen führt und ein inklusiver Unterricht mit höheren Bildungschancen einhergeht (Artiles et al., 2006).

Normative Begründungen für inklusive Bemühungen beziehen sich auf die Frage, wie sich Menschen und Gemeinschaften moralisch richtig verhalten sollten (Dederich, 2013; Horster, 2015). Inklusive Bemühungen lassen sich zum Beispiel auch mit Hegels Anerkennungsparadigma begründen (Horster, 2015). Hegel (zitiert nach Horster, 2015) geht dabei von einer Gleichheit der Verschiedenheit aus (s. a. Prinzip der egalitären Differenz; Prengel, 2013): alle Menschen gleichen sich darin, dass sie sich von anderen unterscheiden. Damit ist eine gegenseitige Anerkennung verbunden, ohne die ein gesellschaftliches Zusammenleben nicht möglich wäre (Horster, 2015). Das normative Element nach Hegel bestehe Horster zufolge darin, an dieser Anerkennung festzuhalten. Somit kann die Wertschätzung und Akzeptanz jeglicher Diversität als normatives Prinzip der Inklusion verstanden werden (Dederich, 2013; Sauter, 2013). Diese normative Begründung fokussiert auf die Anerkennung und Wertschätzung von Heterogenität als ein Ziel von inklusiver Beschulung anstatt auf die effektive Förderung von Schülerinnen und Schülern mit SPF.

Empirische Befunde können normative Begründungen für inklusive Bemühungen im Bildungssystem unterstützen, sie können aber auch in Widerspruch zu normativen Argumenten stehen. Werden zum Beispiel Vorteile leistungshomogener Lerngruppen auf die Leistungsentwicklung von Schülerinnen und Schülern mit SPF berichtet (McLeskey & Waldron, 2011), stünden diese Ergebnisse in Widerspruch zu normativen Prinzipien der Wertschätzung und Pädagogik der Vielfalt (Prengel, 2013). Hier stellt sich die Frage, wie mit solchen Widersprüchen umgegangen werden sollte. Auf der einen Seite wird argumentiert, dass die Partizipation von Kindern und Jugendlichen mit SPF am gemeinsamen Unterricht der allgemeinen Schule ein Menschenrecht ist und keiner zusätzlichen empirischen Begründung beziehungsweise Berechtigung bedarf (Dederich, 2013; Haug, 2014; Moser, 2012). Auf der anderen Seite wird die gemeinsame normative und empirische Betrachtung von Inklusion als bereichernd angesehen (Artiles et al., 2006; Dyson, 2014; Göransson & Nilholm, 2014b; Sauter, 2013). Auch wenn die Idee der Inklusion aus normativer Sicht nicht empirisch überprüft werden kann, so sind deren Umsetzungen sowie die Folgen und Effekte empirisch erforschbar (Dederich, 2013; Veber, 2016). Auch in der UN-Behindertenrechtskonvention (Artikel 31, 2008) wird darauf hingewiesen, dass empirische Daten bereitgestellt werden müssen, um die Entwicklung eines inklusiven Bildungssystems evaluieren zu können. Dabei können empirische Studien zusätzliche Handlungsimpulse für die Umsetzung von Inklusion

bereitstellen. Die Prinzipien und Grundsätze inklusiver Pädagogik werden zwar normativ abgeleitet und begründet (Veber, 2016), sollten aber auch empirisch abgesichert werden. Studien, die methodisch sauber zeigen können, wie Schulen und Klassen inklusiver werden und dass diese Veränderungen mit Vorteilen für die schulische Entwicklung der Kinder verbunden sind, hätten eine stärkere Überzeugungskraft als normative Argumente allein (Göransson & Nilholm, 2014b).

5.5.2 Implikationen für die pädagogische Praxis

Sollen aus den Ergebnissen dieser Dissertation Implikationen für die pädagogische Praxis abgeleitet werden, ergeben sich zwei Fragen. Zum einen stellt sich die Frage, wie der Unterricht in inklusiven Klassen so gestaltet werden kann, dass alle Schülerinnen und Schüler Leistungsfortschritte erzielen. Zum anderen werden für die pädagogische Praxis Hinweise benötigt, wie auch in allgemeinen Schulen eine hohe schulische Motivation von Schülerinnen und Schülern mit SPF erreicht werden kann.

Handlungsempfehlungen zur Unterrichtsgestaltung in inklusiven Schulen lassen sich aus Forschungssynthesen und Beispielen erfolgreicher Schulen ableiten (Ainscow et al., 2005; Booth & Ainscow, 2002, 2011; Dyson et al., 2002; Forlin et al., 2013; Gersten et al., 2009). Prengel (2013) fasst im Rahmen einer wissenschaftlichen Expertise zentrale Merkmale des inklusiven Unterrichts in der Primarstufe zusammen. Eine Voraussetzung für den Einbezug von Schülerinnen und Schülern mit SPF in allgemeine Schulen ist, dass sonderpädagogisch ausgebildete Pädagoginnen und Pädagogen flexibel zur Verfügung stehen (Prengel, 2013). Sie sollten Schülerinnen und Schüler mit Lernschwierigkeiten unabhängig von der Feststellung eines SPF präventiv fördern und mit regulären Lehrkräften sowie zusätzlichen Fachkräften, wie zum Beispiel Sozialpädagoginnen und -pädagogen, kooperieren (Booth & Ainscow, 2002). Das gemeinsame Lehren (*Co-Teaching*; Scruggs, Mastropieri & McDuffie, 2007) durch sonderpädagogische und allgemeine Lehrkräfte kann dabei unterschiedlich ausgestaltet werden. Neben einer zusätzlichen Förderung einzelner Schülerinnen und Schüler durch Sonderpädagoginnen und Sonderpädagogen kann zum Beispiel die Klasse in verschiedene Lerngruppen aufgeteilt werden oder mehrere Lehrkräfte unterrichten eine Klasse gemeinsam.

Weiterhin schlägt Prengel (2013) vor, im Unterricht anspruchsvolle Lehrpläne einzusetzen, die sich an den Bildungsstandards orientieren und hohe Leistungserwartungen an alle

Schülerinnen und Schüler formulieren (s. a. Eccles, 2004). Allerdings sollten diese Lehrpläne individualisiert werden können, um auf die Heterogenität der Schülerschaft angemessen einzugehen. Die Anpassung des Lernmaterials und der Unterrichtsmethoden an die Lernvoraussetzungen einzelner Schülerinnen und Schüler erfordert eine hohe Strukturierung und Vorbereitung durch die Lehrkräfte (Scruggs et al., 2012). Das stellt die Lehrkräfte vor besondere Herausforderungen bei der Unterrichtsgestaltung, auf die in Fort- und Weiterbildungen eingegangen werden könnte (Prenzel, 2013). Insbesondere sollten die Lehrkräfte im Einsatz von diagnostischen Verfahren (Prozessdiagnostik) geschult werden, um passende Lernziele und -aufgaben für ihre Schülerinnen und Schüler auszuwählen (Tomlinson, 2015). Der Einsatz individueller Lernaufgaben sollte zudem an allgemeine Lehrpläne anknüpfen, um kooperatives Lernen an gemeinsamen Lernzielen zwischen Schülerinnen und Schülern einer Klasse zu fördern. Es sollte also eine ausgewogene Balance zwischen Individualisierung durch differenzierende Unterrichtselemente und gemeinsamen Lernaufgaben und -ziele gefunden werden (Prenzel, 2013). Eine zentrale Herausforderung besteht somit in der individuellen Ausrichtung des Unterrichts einerseits und der Betonung von Gemeinsamkeiten der Lerngruppe andererseits (Grosche, 2015; Horster, 2015).

Zusätzlich zu diesen Empfehlungen zur Organisation des Unterrichts in inklusiven Klassen wurden konkrete Unterrichtsmethoden zur Leistungsförderung von Schülerinnen und Schülern in den Blick genommen (Gersten et al., 2009). Als besonders wirksam erwiesen sich in Interventionsstudien explizite Instruktionen (Gersten et al., 2009). Explizite Instruktionen beinhalten eine schrittweise Anleitung zur Entwicklung des Lösungsweges durch die Lehrkräfte, der auf spezifische Lernaufgaben ausgerichtet ist. Die Schülerinnen und Schüler wenden diesen Lösungsweg anschließend auf sehr ähnliche Lernaufgaben an (Gersten et al., 2009). Zusätzlich zu einer engen Anleitung der Schülerinnen und Schüler bei spezifischen Lernaufgaben war auch die Vermittlung allgemeiner Strategien zur Lösung von Lernaufgaben hilfreich (Gersten et al., 2009). Solche Strategien umfassen zum Beispiel die Hervorhebung zentraler Begriffe in einer Aufgabe, die für die Lösung wichtig sind, sowie die Aufteilung der Aufgabe in verschiedene Schritte und das Überprüfen des gewählten Lösungsweges. Sowohl die Effektivität expliziter Instruktionen als auch der Vermittlung allgemeiner Strategien zur Bearbeitung von Lernaufgaben weisen darauf hin, dass Kinder und Jugendliche mit Lernbeeinträchtigung von einer hohen Strukturierung und engen Begleitung durch Lehrkräfte bei der Bearbeitung von Lernaufgaben profitieren (Tillmann & Wischer, 2006).

Im Gegensatz dazu zeigten sich in der Metaanalyse von Gersten et al. (2009) insgesamt keine bedeutsamen Effekte von Peer-Tutoring auf die mathematischen Leistungen von Schülerinnen und Schülern mit Lernbeeinträchtigungen. Diese Ergebnisse weisen darauf hin, dass das gemeinsame Lernen von Schülerinnen und Schülern mit und ohne SPF stark vorstrukturiert werden sollte. Außerdem sollten die Leistungsstände der Kinder, die gemeinsam eine Lernaufgabe bearbeiten, nicht zu weit auseinander liegen. So gibt es Hinweise darauf, dass leistungstärkere Schülerinnen und Schüler mit SPF, die ähnliche schulische Kompetenzen wie ihre Mitschülerinnen und Mitschüler aufweisen, eher vom Peer-Tutoring profitieren (Gersten et al., 2009). Dieses Ergebnis relativiert die Ergebnisse der dritten Teilstudie, in der positive Zusammenhänge zwischen dem mittleren Fähigkeitsniveau einer Klasse und individuellen schulischen Kompetenzen von Schülerinnen und Schülern mit SPF festgestellt wurden. Es scheint nicht auszureichen, alle Kinder und Jugendlichen mit SPF in leistungstärkeren Lerngruppen zu unterrichten. Die Bildung von Lerngruppen, die aus Kindern und Jugendlichen mit und ohne SPF bestehen, muss vielmehr gut vorbereitet und strukturiert werden.

Zusammenfassend lässt sich festhalten, dass sich für Kinder und Jugendliche mit SPF und Schülerinnen und Schülern ohne SPF ähnliche Unterrichtsmethoden als effektiv erwiesen haben (Rix & Sheehy, 2014). Deshalb argumentieren Rix und Sheehy, dass sich eine inklusive Pädagogik durch die gleichen Prinzipien auszeichnet wie ein guter Unterricht für Schülerinnen und Schüler ohne SPF. Schülerinnen und Schüler mit einem SPF in den Bereichen *Lernen*, *Sprache* und *emotionale und soziale Entwicklung* scheinen von ähnlichen Unterrichtsmethoden zu profitieren wie andere leistungsschwächere Schülerinnen und Schüler, zum Beispiel von einer hohen Strukturierung der Lernaufgaben und einer stärkeren Anleitung durch die Lehrkräfte.

Neben den schulischen Kompetenzen sollte in inklusiven Klassen auch die schulische Motivation der Schülerinnen und Schüler mit SPF gefördert werden (Prengel, 2013). Insbesondere bei Schülerinnen und Schülern mit SPF kann eine geringe schulische Motivation dazu beitragen, dass Lernschwierigkeiten persistieren (Deci & Chandler, 1986). Eine Grundlage für die Förderung der schulischen Motivation ist ein unterstützendes Klassenklima mit wertschätzenden und unterstützenden Beziehungen zwischen Lehrkräften und Schülerinnen und Schülern sowie der Mitschülerinnen und Mitschüler untereinander (Deci & Chandler, 1986; Prengel, 2013). Auch in Teilstudie 4 zeigten sich positive Zusammenhänge zwischen der schulischen Motivation von Kindern mit SPF und der wahrgenommenen emotionalen Unterstützung durch Mitschülerinnen und Mitschüler. Schülerinnen und Schüler mit SPF

könnten darüber hinaus von einer geringeren Wettbewerbsorientierung im Unterricht profitieren, in dem die Beiträge jeder Schülerin und jedes Schülers wertgeschätzt werden (Dijkstra et al., 2008; Marsh et al., 2008). Um den negativen Zusammenhang zwischen der mittleren Leistungsstärke der Klasse und der schulischen Motivation von Schülerinnen und Schülern mit SPF abzuschwächen, sollten in inklusiven Klassen verstärkt individuelle Leistungsrückmeldungen eingesetzt werden (Dijkstra et al., 2008). So könnten Leistungsbeurteilungen in Form von Schulnoten mit Rückmeldungen auf Grundlage einer individuellen Bezugsnorm kombiniert und auf individuelle Stärken und Schwächen ausgerichtet werden (Deci & Chandler, 1986; Marsh et al., 2008).

5.6 Fazit

Schülerinnen und Schüler mit SPF werden zunehmend in allgemeinen Schulen gemeinsam mit Schülerinnen und Schülern ohne SPF anstatt in Förderschulen unterrichtet (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2016). Es besteht aber noch Forschungsbedarf, wie die Beschulung in Förderschulen oder allgemeinen Schulen mit der schulischen Entwicklung von Schülerinnen und Schülern mit SPF zusammenhängt. Die vorliegende Dissertation trägt zum Forschungsstand bei, indem Zusammenhänge der Beschulungsart mit den schulischen Kompetenzen und der schulischen Motivation von Kindern mit SPF am Ende der Grundschule untersucht wurden. Dabei zeigte sich, dass Kinder mit SPF in Grundschulen höhere schulische Kompetenzen in Deutsch und Mathematik erzielten, aber ein geringeres akademisches Selbstkonzept und eine niedrigere Lernfreude berichteten als vergleichbare Kinder mit SPF in Förderschulen. Die Arbeit erweitert den Forschungsstand um eine nach Förderschwerpunkt differenzierte Betrachtung und um Ergebnisse zur Klassenzusammensetzung als mögliche Erklärung für Schulartunterschiede.

In dieser Dissertation liegt der Fokus auf einem spezifischen Aspekt von Inklusion: die gemeinsame Beschulung von Schülerinnen und Schülern mit SPF in allgemeinen Schulen anstatt einer separaten Beschulung in Förderschulen (Ruijs & Peetsma, 2009). Die Platzierung von Schülerinnen und Schülern mit SPF in Förderschulen oder allgemeinen Schulen zählt zu den am häufigsten untersuchten Merkmalen von Inklusion (Artiles et al., 2006; Dyson et al., 2002; Ruijs & Peetsma, 2009). Ein Vorteil des hier gewählten Untersuchungsansatzes ist auch, dass sich dieses Merkmal von Inklusion in Schulleistungsstudien gut operationalisieren

lässt. So werden in vielen Schulleistungsstudien Informationen zum sonderpädagogischen Förderbedarf und der besuchten Schulart der Schülerinnen und Schüler erhoben (z. B. Pant et al., 2013; Prenzel, Sälzer, Klieme & Köller, 2013; Stanat, Böhme, Schipolowski & Haag, 2016; Stanat et al., 2012).

Allerdings wird die Operationalisierung von Inklusion über die Beschulungsart als zu reduktionistisch kritisiert (Florian, 2014). Die Platzierung von Kindern und Jugendlichen mit SPF in allgemeinen Schulen stelle demnach einen untergeordneten Aspekt von Inklusion dar (Göransson & Nilholm, 2014a). Wichtig sei vielmehr eine Betrachtung der konkreten Ausgestaltung des Unterrichts und die soziale Partizipation der Schülerinnen und Schüler mit SPF im gemeinsamen Unterricht (Florian, 2014; Werning, 2010). Zusätzlich sollten unter einem weiten Verständnis von Inklusion weitere Schülergruppen in den Blick genommen werden, die im Schulsystem von Marginalisierung bedroht sind (z. B. Kinder und Jugendliche mit Teilleistungsstörungen oder geringen Kenntnissen der Verkehrssprache). Dabei könnte auch darauf verzichtet werden, einzelne Merkmalsausprägungen von Schülerinnen und Schülern zu untersuchen und stattdessen eine kontinuierliche Betrachtung verschiedener Heterogenitätsdimensionen vorgenommen werden (Biewer, 2010; Budesco & Budesco, 2012; Hinz, 2002). Außerdem könnten zukünftige Studien intersektional ausgerichtet werden und verschiedene Dimensionen von Heterogenität in ihrem Zusammenspiel untersuchen (Emmerich & Hormel, 2013; DGfE, 2017).

Eine mögliche Operationalisierung von Inklusion in zukünftigen Studien könnte die Beschreibung des Umgangs mit exklusiven Praktiken sein (Emmerich, 2016). Dabei werden Barrieren für die Partizipation aller Schülerinnen und Schüler am gemeinsamen Unterricht analysiert und anschließend Lösungsansätze zu deren Überwindung aufgezeigt (DGfE, 2017; Emmerich, 2016). Diese Perspektive bietet laut Emmerich (2016) einen höheren Forschungsgehalt. So verschöbe sich die Fragestellung dahin, wie sich gesellschaftliche Ungleichheiten und exkludierende Praktiken auf individuelle Bildungsprozesse auswirken. Dieser Ansatz hätte den Vorteil, dass realistischere Erwartungshaltungen zur Wirkung inklusiver Ansätze beim Abbau von Bildungsungleichheiten gebildet werden, da die Komplexität des bestehenden Bildungssystems angemessen berücksichtigt wird (Emmerich, 2016). Außerdem ist dieser Ansatz kompatibel mit einem Verständnis, das Inklusion nicht als einen abgeschlossenen Zustand betrachtet, sondern als einen graduellen Wandel hin zu einem inklusiveren Bildungssystem (Göransson & Nilholm, 2014a).

Literaturverzeichnis

- Abedi, J., Leon, S., Kao, J., Bayley, R., Ewers, N., Herman, J. & Mundhenk, K. (2011). *Accessible reading assessments for students with disabilities: The role of cognitive, grammatical, lexical, and textual/visual features. CRESST Report 785*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Agirdag, O., Houtte, M. V. & Avermaet, P. V. (2012). Why does the ethnic and socio-economic composition of schools influence math achievement? The role of sense of futility and futility culture. *European Sociological Review*, 28, 366–378. <https://doi.org/10.1093/esr/jcq070>
- Ahrbeck, B. (2014). Schulische Inklusion – Möglichkeiten, Dilemmata und Widersprüche. *Soziale Passagen*, 6(1), 5–19. <https://doi.org/10.1007/s12592-014-0154-x>
- Ahrbeck, B. (2016). *Inklusion: Eine Kritik* (3. Aufl.). Stuttgart: Kohlhammer Verlag.
- Ainscow, M. (1991). Effective schools for all: an alternative approach to special needs in education. *Cambridge Journal of Education*, 21, 293–308. <https://doi.org/10.1080/0305764910210304>
- Ainscow, M., Booth, T. & Dyson, A. (2006). *Improving schools, developing inclusion*. London: Routledge.
- Allodi, M. W. (2000). Self-concept in children receiving special support at school. *European Journal of Special Needs Education*, 15, 69–78. <https://doi.org/10.1080/088562500361718>
- Altermatt, E. R. & Pomerantz, E. M. (2005). The implications of having high-achieving versus low-achieving friends: A longitudinal analysis. *Social Development*, 14, 61–81. <https://doi.org/10.1111/j.1467-9507.2005.00291.x>
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)* (5. Aufl.). Washington, DC: American Psychiatric Pub.
- Anderman, E. M. (2002). School effects on psychological outcomes during adolescence. *Journal of Educational Psychology*, 94, 795–809. <https://doi.org/http://dx.doi.org/10.1037/0022-0663.94.4.795>
- Artiles, A. J., Kozleski, E. B., Dorn, S. & Christensen, C. (2006). Learning in inclusive education research: Re-mediating theory and methods with a transformative agenda. *Review of Research in Education*, 30, 65–108. <https://doi.org/10.3102/0091732X030001065>
- Asseburg, R. & Frey, A. (2013). Too hard, too easy, or just right? The relationship between effort or boredom and ability-difficulty fit. *Psychological Test and Assessment Modeling*, 55, 92–104.
- Auernheimer, G. (Hrsg.). (2010). *Schief lagen im Bildungssystem. Die Benachteiligung der Migrantenkinder*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Autorengruppe Bildungsberichtserstattung. (2010). *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demografischen Wandel*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung. (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development*, 59, 87–96. <https://doi.org/10.2307/1130391>
- Baker, E. T., Wang, M. C. & Walberg, H. J. (1994). The effects of inclusion on learning. *Educational Leadership*, 52(4), 33–35.

- Bakker, J. T. A. & Bosman, A. M. T. (2003). Self-image and peer acceptance of Dutch students in regular and special education. *Learning Disability Quarterly*, 26, 5–14. <https://doi.org/10.2307/1593680>
- Bakker, J. T. A., Denessen, E., Bosman, A. M. T., Krijger, E.-M. & Bouts, L. (2007). Sociometric status and self-image of children with specific and general learning disabilities in Dutch general and special education classes. *Learning Disability Quarterly*, 30, 47–62. <https://doi.org/10.2307/30035515>
- Bandura, A. (1971). *Social learning theory*. Morristown, N.J.: General Learning Press.
- Banerji, M. & Dailey, R. A. (1995). A study of the effects of an inclusion model on students with specific learning disabilities. *Journal of Learning Disabilities*, 28, 511–522. <https://doi.org/10.1177/002221949502800806>
- Barth, K. (2016). *Die Diagnostischen Einschätzskalen (DES) zur Beurteilung des Entwicklungsstandes und der Schulfähigkeit: Handanweisung – Aufgabenteil – Auswertungs- und Einschätzbogen – Entwicklungsprofilbogen* (7. Aufl.). München: Ernst Reinhardt Verlag.
- Baumeister, R. F. & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Baumert, J. (2016a). Bildungsgerechtigkeit in Deutschland – ein Überblick. In D. Döring (Hrsg.), *Auf der Suche nach der richtigen Ordnung: Königsteiner Forum 2014* (S. 63–98). Frankfurt a.M.: Societäts-Verlag.
- Baumert, J. (2016b). Leistungen, Leistungsfähigkeit und Leistungsgrenzen der empirischen Bildungsforschung. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker* (S. 215–253). Wiesbaden: Springer VS.
- Baumert, J. & Köller, O. (1998). Nationale und internationale Schulleistungsstudien: was können sie leisten, wo sind ihre Grenzen? *Pädagogik*, 50(6), 12–18.
- Baumert, J., Köller, O. & Schnabel, K.-U. (2000). Schulformen als differentielle Entwicklungsmilieus - Eine ungehörige Fragestellung? In Gewerkschaft Erziehung und Wissenschaft (Hrsg.), *Messung sozialer Motivation : Eine Kontroverse* (S. 28–69). Frankfurt/M.: Bildungs- und Förderungswerk der GEW im DGB.
- Baumert, J., Maaz, K., Stanat, P. & Watermann, R. (2009). Schulkomposition oder Institution - was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule*, 101, 33–46.
- Baumert, J., Nagy, G. & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, 83, 1347–1367. <https://doi.org/10.1111/j.1467-8624.2012.01779.x>
- Baumert, J., Roeder, P. M., Sang, F. & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. *Zeitschrift für Pädagogik*, 5, 639–660.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–407). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bear, G. G., Minke, K. M. & Manning, M. A. (2002). Self-concept of students with learning disabilities: A meta-analysis. *School Psychology Review*, 31, 405–427.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik: Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie*, 20, 233–242. <https://doi.org/10.1024/1010-0652.20.4.233>
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104, 682–699. <https://doi.org/10.1037/a0027608>

- Becker, R. (2007). Soziale Ungleichheit von Bildungschancen und Chancengerechtigkeit. In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (2. Aufl., S. 157–185). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Beddow, P. A. (2011). Beyond universal design. Accessibility theory to advance testing for all students. In M. K. Russell & M. Kavanaugh (Hrsg.), *Assessing students in the margin: Challenges, strategies, and techniques* (S. 381–406). Charlotte, NC: Information Age Publishing.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Biermann, J. & Powell, J. J. W. (2014). Institutionelle Dimensionen inklusiver Schulbildung – Herausforderungen der UN-Behindertenrechtskonvention für Deutschland, Island und Schweden im Vergleich. *Zeitschrift für Erziehungswissenschaft*, 17, 679–700.
<https://doi.org/https://doi.org/10.1007/s11618-014-0588-0>
- Biewer, G. (2010). *Grundlagen der Heilpädagogik und inklusiven Pädagogik*. Bad Heilbrunn: Klinkhardt.
- Biewer, G. (2012). Die neue Welt der Bildungsstandards und ihre erziehungswissenschaftliche Rezeption aus der Perspektive einer Inklusiven Pädagogik. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 1, 9–21. <https://doi.org/http://dx.doi.org/10.2378/vhn2012.art01d>
- Blanton, H., Buunk, A. P., Gibbons, F. X. & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, 76, 420–430.
<https://doi.org/10.1037/0022-3514.76.3.420>
- Bleidick, U. (1988). *Betrifft Integration: behinderte Schüler in allgemeinen Schulen. Konzepte der Integration: Darstellung und Ideologiekritik*. Berlin: Carl Marhold.
- Bleidick, U., Rath, W. & Schuck, K. D. (1995). Die Empfehlungen der Kultusministerkonferenz zur sonderpädagogischen Förderung in den Schulen der Bundesrepublik Deutschland. *Zeitschrift für Pädagogik*, 41, 247–264.
- Böhme, K., Kohrt, P., Haag, N. & Weirich, S. (2017). Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 369–385). Münster: Waxmann.
- Booth, T. & Ainscow, M. (2002). *Index for inclusion: Developing learning and participation in schools*. Bristol: Centre for Studies on Inclusive Education (CSIE).
- Booth, T. & Ainscow, M. (2011). *Index for inclusion: Developing learning and participation in schools*. Bristol: Centre for Studies on Inclusive Education (CSIE).
- Bos, W., Müller, S. & Stubbe, T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society*. New York: John Wiley & Sons.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Brodesser, E. (2016). *Kinder aus zugewanderten Familien in Förderschulen: Vertiefende Auswertung der Berliner Erhebung „BELLA“ zu herkunftsbezogenen Disparitäten*. Wiesbaden: Springer VS.
- Brophy, J. (1983). Conceptualizing student motivation. *Educational Psychologist*, 18, 200–215.
<https://doi.org/10.1080/00461528309529274>

- Budescu, D. V. & Budescu, M. (2012). How to measure diversity when you must. *Psychological Methods*, 17, 215–227. <https://doi.org/10.1037/a0027129>
- Butler, R. & Marinov-Glassman, D. (1994). The effects of educational placement and grade level on the self-perceptions of low achievers and students with learning disabilities. *Journal of Learning Disabilities*, 27, 325–334. <https://doi.org/10.1177/002221949402700509>
- Büttner, G., Dacheneder, W., Schneider, W. & Weyer, K. (2008). *Frostigs Entwicklungstest der visuellen Wahrnehmung – 2 (FEW-2)*. Göttingen: Hogrefe.
- Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation*. Washington, DC: American Psychological Association.
- Carlberg, C. & Kavale, K. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *The Journal of Special Education*, 14, 295–309. <https://doi.org/10.1177/002246698001400304>
- Carlisle, J. F., Cortina, K. S. & Zeng, J. (2010). Reading achievement in Reading First schools in Michigan. *Journal of Literacy Research*, 42, 49–70. <https://doi.org/10.1080/10862960903583236>
- Chapman, J. W. (1988). Learning disabled children's self-concepts. *Review of Educational Research*, 58, 347–371. <https://doi.org/10.3102/00346543058003347>
- Chiu, C. W. T. & Pearson, P. D. (1999, Juni). *Synthesizing the effects of test accommodations for special education and limited English proficient students*. Konferenzbeitrag gehalten auf der Annual Large-scale Assessment Conference of the Council of Chief State School Officers, Snowbird, UT. Konferenzbeitrag abgerufen von <https://eric.ed.gov/?id=ED433362>
- Chmielewski, A. K., Dumont, H. & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, 50, 925–957. <https://doi.org/10.3102/0002831213489843>
- Chudowsky, N. & Pellegrino, J. W. (2003). Large-scale assessments that support learning: What will it take? *Theory Into Practice*, 42, 75–83. https://doi.org/10.1207/s15430421tip4201_10
- Cialdini, R., Borden, R., Thorne, A., Walker, M. R., Freeman, S. & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, 34, 366–375. <https://doi.org/10.1037/0022-3514.34.3.366>
- Cole, C. M., Waldron, N. & Majd, M. (2004). Academic progress of students across inclusive and traditional settings. *Mental Retardation*, 42, 136–144.
- Cole, N. S. & Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38, 369–382.
- Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 119, 51–69. <https://doi.org/10.1037/0033-2909.119.1.51>
- Connell, J. P. & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Hrsg.), *Self processes and development* (S. 43–77). Hillsdale, NJ, England: Lawrence Erlbaum.
- Cook, L., Eignor, D., Steinberg, J., Sawaki, Y. & Cline, F. (2009). Using factor analysis to investigate the impact of accommodations on the scores of students with disabilities on a reading comprehension assessment. *Journal of Applied Testing Technology*, 10(2), 1–33.
- Cooney, G., Jahoda, A., Gumley, A. & Knott, F. (2006). Young people with intellectual disabilities attending mainstream and segregated schooling: perceived stigma, social comparison and future aspirations. *Journal of Intellectual Disability Research*, 50, 432–444. <https://doi.org/10.1111/j.1365-2788.2006.00789.x>
- Corcoran, K., Crusius, J. & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Hrsg.), *Theories in social psychology* (S. 119–139). Oxford, UK: Wiley-Blackwell.
- Cormier, D. C., Altman, J., Shyyan, V. & Thurlow, M. L. (2010). *A summary of the research on the effects of test accommodations: 2007-2008 (Technical Report 56)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Crabtree, J. W. & Meredith, C. (2000). Self-concept and social comparisons in learning disabled students attending mainstream and special schools: Does integration have an impact? In R. G. Craven & H. W. Marsh (Hrsg.), *Self-concept theory, research and practice: Advances for the new millennium* (S. 187–193). Sydney: SELF Research Centre, University of Western Sydney.

- Creemers, B. P. M. & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, 7, 197–228. <https://doi.org/10.1080/0924345960070301>
- Dai, D. Y. & Rinn, A. N. (2008). The big-fish-little-pond effect: What do we know and where do we go from here? *Educational Psychology Review*, 20, 283–317. <https://doi.org/10.1007/s10648-008-9071-x>
- Deci, E. L. & Chandler, C. L. (1986). The importance of motivation for the future of the LD field. *Journal of Learning Disabilities*, 19, 587–594.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G. & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, 26, 325–346. <https://doi.org/10.1080/00461520.1991.9653137>
- Dederich, M. (2013). *Ethische Aspekte der Inklusion*. Zugriff am 05.09.2017 unter http://www.inklusion-lexikon.de/Ethik_Dederich.pdf
- Demaray, M. K. & Malecki, C. K. (2002). The relationship between perceived social support and maladjustment for students at risk. *Psychology in the Schools*, 39, 305–316. <https://doi.org/10.1002/pits.10018>
- DGfE = Deutsche Gesellschaft für Erziehungswissenschaft. (2017). *Inklusion: Bedeutung und Aufgabe für die Erziehungswissenschaft*. Zugriff am 30.05.2017 unter http://www.dgfe.de/fileadmin/OrdnerRedakteure/Stellungnahmen/2017.01_Inklusion_Stellungnahmen.pdf
- Diamond, J. B., Randolph, A. & Spillane, J. P. (2004). Teachers' expectations and sense of responsibility for student learning: The importance of race, class, and organizational habitus. *Anthropology & Education Quarterly*, 35, 75–98. <https://doi.org/10.1525/aeq.2004.35.1.75>
- Dietze, T. (2011). Sonderpädagogische Förderung in Zahlen - Ergebnisse der Schulstatistik 2009/10 mit einem Schwerpunkt auf der Analyse regionaler Disparitäten. *Zeitschrift für Inklusion*, 6(2), 1–21.
- Dietze, T. (2012). Zum Stand der sonderpädagogischen Förderung in Deutschland. Die Schulstatistik 2010/11. *Zeitschrift für Heilpädagogik*, 63, 26–31.
- Dijkstra, P., Kuiper, H., Werf, G. van der, Buunk, A. P. & Zee, Y. G. van der. (2008). Social comparison in the classroom: A review. *Review of Educational Research*, 78, 828–879. <https://doi.org/10.3102/0034654308321210>
- Ditton, H. (2007). Schulübertritte, Geschlecht und soziale Herkunft. In H. Ditton (Hrsg.), *Kompetenzaufbau und Laufbahnen im Schulsystem. Ergebnisse einer Längsschnittuntersuchung an Grundschulen* (S. 53–87). Münster: Waxmann.
- Ditton, H., Krüskens, J. & Schauenberg, M. (2005). Bildungsungleichheit — der Beitrag von Familie und Schule. *Zeitschrift für Erziehungswissenschaft*, 8, 285–304. <https://doi.org/10.1007/s11618-005-0138-x>
- Dreeben, R. & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, 61, 129–142. <https://doi.org/10.2307/2112622>
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183. <https://doi.org/10.2378/peu2013.art14d>
- Duncan, C., Jones, K. & Moon, G. (1998). Context, composition and heterogeneity: using multilevel models in health research. *Social Science & Medicine*, 46, 97–117.
- Duncan, G. J., Magnuson, K. A. & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1, 59–80. <https://doi.org/10.1080/15427609.2004.9683330>
- Dunn, L. M. (1968). Special education for the mildly retarded-is much of it justifiable? *Exceptional Children*, 35, 5–22.
- Dyson, A. (2014). A response to Göransson and Nilholm. *European Journal of Special Needs Education*, 29, 281–282. <https://doi.org/10.1080/08856257.2014.933542>
- Eberwein, H. (1996). *Handbuch Lernen und Lern-Behinderungen: Aneignungsprobleme-Neues Verständnis von Lernen-Integrationspädagogische Lösungsansätze* (1. Aufl.). Weinheim: Beltz.

- Eccles, J. S. (2004). Schools, academic motivation, and stage-environment. In R. M. Lerner & L. Steinberg (Hrsg.), *Handbook of Adolescent Psychology* (S. 125–153). Hoboken, NJ: John Wiley & Sons.
- Eccles, J. S. & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225–241. <https://doi.org/10.1111/j.1532-7795.2010.00725.x>
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Eccles (Parsons), J., Adler, T. E., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives. Psychological and sociological approaches* (S. 75–146). San Francisco, CA: W. H. Freeman.
- Elbaum, B. (2002). The self-concept of students with learning disabilities: A meta-analysis of comparisons across different placements. *Learning Disabilities Research & Practice*, 17, 216–226. <https://doi.org/10.1111/1540-5826.00047>
- Ellger-Rüttgardt, S. (2004). Sonderpädagogik - ein blinder Fleck der Allgemeinen Pädagogik? Eine Replik auf den Aufsatz von Dagmar Hänsel. *Zeitschrift für Pädagogik*, 50, 416–429.
- Elliott, S. N., Beddow, P. A., Kurz, A. & Kettler, R. J. (2011). Creating access to instruction and tests of achievement: Challenges and solutions. In S. N. Elliott, R. J. Kettler, P. A. Beddow & A. Kurz (Hrsg.), *Handbook of accessible achievement tests for all students* (S. 1–16). New York, NY: Springer.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Erlbaum: Hillsdale, NJ.
- Emmerich, M. (2016). Differenz und Differenzierung im Bildungssystem: Schulische Grammatik der Inklusion/Exklusion. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 42–57). Weinheim: Beltz Juventa.
- Emmerich, M. & Hormel, U. (2013). *Heterogenität - Diversity - Intersektionalität*. Wiesbaden: Springer VS.
- Ennis, C. D. (1992). Reconceptualizing learning as a dynamical system. *Journal of Curriculum and Supervision*, 7, 115–130.
- Europäische Kommission. (1996). *Die Charta von Luxemburg*. Brüssel: Europäische Kommission.
- European Agency for Development in Special Needs Education. (2011). *Mapping the Implementation of Policy for Inclusive Education: An exploration of challenges and opportunities for developing indicators*. Odense, Denmark: European Agency for Development in Special Needs Education.
- European Agency for Development in Special Needs Education. (2012). *Special Needs Education Country Data 2012*. Odense, Denmark: European Agency for Development in Special Needs Education.
- Farrell, P., Dyson, A., Polat, F., Hutcheson, G. & Gallannaugh, F. (2007). The relationship between inclusion and academic achievement in English mainstream schools. *School Effectiveness and School Improvement*, 18, 335–352. <https://doi.org/10.1080/09243450701442746>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. <https://doi.org/10.1177/001872675400700202>
- Fischer, E., Heimlich, U., Kahlert, J. & Lelgemann, R. (2014). *Bericht zum I. Beauftragungszeitraum des Wissenschaftlichen Beirats „Inklusion“*. München und Würzburg: Wissenschaftlicher Beirat „Inklusion“ beauftragt durch den Bayerischen Landtag.
- Florian, L. (2014). What counts as evidence of inclusive education? *European Journal of Special Needs Education*, 29, 286–294. <https://doi.org/10.1080/08856257.2014.933551>
- Forlin, C., Chambers, D., Loreman, T., Deppler, J. & Sharma, U. (2013). *Inclusive education for students with disability: A review of the best evidence in relation to theory and practice*. The Australian Research Alliance for Children and Youth. Zugriff am 27.10.2016 unter http://researchonline.nd.edu.au/edu_article/141
- Forman, E. A. (1988). The effects of social support and school placement on the self-concept of LD students. *Learning Disability Quarterly*, 11, 115–124. <https://doi.org/10.2307/1510989>
- Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology*, 46, 1454–1480. <https://doi.org/10.1037/a0020204>

- Frank, A., Kirschhock, E.-M. & Martschinke, S. (2003). Wie erkenne ich „Risikokinder“ im Schriftspracherwerb? Diagnose phonologischer Bewusstheit. *Grundschule*, 35(9), 58–59.
- Frey, D., Dauenhauer, D., Parge, O., Haisch, J., Frey, D. & Irle, M. (1993). Die Theorie sozialer Vergleichsprozesse. In *Theorien der Sozialpsychologie, Band I: Kognitive Theorien* (2. Aufl., S. 81–121). Bern: Hans Huber.
- Fuchs, D., Fuchs, L. S., Mathes, P. G. & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206. <https://doi.org/10.2307/1163346>
- Fuchs, L. S. & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research & Practice*, 16, 174–181. <https://doi.org/10.1111/0938-8982.00018>
- Fulk, B. M., Brigham, F. J. & Lohman, D. A. (1998). Motivation and self-regulation: A comparison of students with learning and behavior problems. *Remedial and Special Education*, 19, 300–309. <https://doi.org/10.1177/074193259801900506>
- Furrer, C. & Skinner, E. (2003). Sense of relatedness as a factor in children’s academic engagement and performance. *Journal of Educational Psychology*, 95, 148–162. <https://doi.org/10.1037/0022-0663.95.1.148>
- Füssel, H.-P. & Kretschmann, R. (1993). *Gemeinsamer Unterricht für behinderte und nichtbehinderte Kinder: pädagogische und juristische Voraussetzungen ; Gutachten im Auftrag der Max-Traeger-Stiftung*. Witterschlick/Bonn: Wehle.
- Gamoran, A. (1986). Instructional and institutional effects of ability grouping. *Sociology of Education*, 59, 185–198. <https://doi.org/10.2307/2112346>
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P. & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79, 1202–1242. <https://doi.org/10.3102/0034654309334431>
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice Hall.
- Goffman, E. (1986). *Stigma: Notes on the management of spoiled identity* (Reissue edition). New York: Touchstone.
- Gomolla, M. (2006). Fördern und Fordern allein genügt nicht! Mechanismen institutioneller Diskriminierung von Migrantenkindern und -jugendlichen im deutschen Schulsystem. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem* (S. 87–102). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gomolla, M. & Radtke, F.-O. (2009). *Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Göransson, K. & Nilholm, C. (2014a). Conceptual diversities and empirical shortcomings – a critical analysis of research on inclusive education. *European Journal of Special Needs Education*, 29, 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Göransson, K. & Nilholm, C. (2014b). A continuing need for conceptual analysis into research on inclusive education: response to commentators. *European Journal of Special Needs Education*, 29, 295–296. <https://doi.org/10.1080/08856257.2014.934035>
- Gorges, J., Neumann, P., Wild, E., Stranghöner, D. & Lütje-Klose, B. (2017, März). *Reziproke Effekte zwischen Fähigkeitsselbstkonzept und Leistung bei Grundschulkindern mit sonderpädagogischem Förderbedarf im Förderschwerpunkt Lernen (SFB-L)*. Konferenzbeitrag gehalten auf der 5. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Heidelberg.
- Gottfried, M. A. (2014). Classmates with disabilities and students’ noncognitive outcomes. *Educational Evaluation and Policy Analysis*, 36, 20–43.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Gresch, C., Piezunka, A. & Solga, H. (2014). Realisierbarkeit einer Ergänzungsstichprobe von Integrationsschülerinnen und -schülern im Rahmen des Nationalen Bildungspanels (NEPS): Möglichkeiten und Perspektiven. In *NEPS Working Paper* (Bd. 37). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

- Grimm, H. & Schöler, H. (Hrsg.). (1991). *Heidelberger Sprachentwicklungstest (HSET)* (2. Aufl.). Göttingen: Hogrefe Verlag.
- Gröblich, C., Scharenberg, K. & Bos, W. (2009). Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? *Journal for Educational Research Online*, 1, 86–105.
- Grolnick, W. S. & Ryan, R. M. (1990). Self-perceptions, motivation, and adjustment in children with learning disabilities: A multiple group comparison study. *Journal of Learning Disabilities*, 23, 177–184.
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.
- Grosche, M. & Vock, M. (im Druck). Inklusion. In D. H. Rost, J. Sparfeldt & S. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. Aufl.). Weinheim: Beltz.
- Haag, N., Böhme, K. & Stanat, P. (2012). Zuwanderungsbezogene Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 209–235). Münster: Waxmann.
- Haeberlin, U. (1991). Die Integration von leistungsschwachen Schülern. Ein Überblick über empirische Forschungsergebnisse zu Wirkungen von Regelklassen, Integrationsklassen und Sonderklassen auf „Lernbehinderte“. *Zeitschrift für Pädagogik*, 37, 167–189.
- Hallinan, M. T. (1988). School composition and learning: A critique of the Dreeben-Barr model. *Sociology of Education*, 61, 143–146.
- Händel, M., Lockl, K., Heydrich, J., Weinert, S. & Artelt, C. (in press). Kompetenztestung bei Schülerinnen und Schülern mit Förderschwerpunkt Lernen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 221–242). Wiesbaden: Springer VS.
- Hänsel, D. (2005). Die Historiographie der Sonderschule. Eine kritische Analyse. *Zeitschrift für Pädagogik*, 51, 101–115.
- Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18, 527–544. <https://doi.org/10.1002/jae.741>
- Hanushek, E. A., Kain, J. F. & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics*, 84, 584–599. <https://doi.org/10.1162/003465302760556431>
- Harker, R. & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15, 177–199. <https://doi.org/10.1076/sesi.15.2.177.30432>
- Harrington, S. A. (1997). Full inclusion for students with learning disabilities: A review of the evidence. *School Community Journal*, 7(1), 63–71.
- Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, 112, 1163–1197.
- Harrison, J. R., Bunford, N., Evans, S. W. & Owens, J. S. (2013). Educational accommodations for students with behavioral challenges: A systematic review of the literature. *Review of Educational Research*, 83, 551–597. <https://doi.org/10.3102/0034654313497517>
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Aufl., S. 135–163). Heidelberg: Springer.
- Hartmann, U., Decristan, J. & Klieme, E. (2016). Unterricht als Feld evidenzbasierter Bildungspraxis? Herausforderungen und Potenziale für einen wechselseitigen Austausch von Wissenschaft und Schulpraxis. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker* (S. 179–199). Wiesbaden: Springer VS.
- Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., ... Zoelch, C. (2012). *Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12)* (1. Aufl.). Göttingen: Hogrefe.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, New York: Routledge.

- Haug, P. (2014). Empirical shortcomings? A comment on Kerstin Göransson and Claes Nilholm, 'Conceptual diversities and empirical shortcomings - a critical analysis of research on inclusive education'. *European Journal of Special Needs Education*, 29, 283–285.
<https://doi.org/10.1080/08856257.2014.933548>
- Heimlich, U., Lutz, S. & Wilfert de Icaza, K. (2013). *Ratgeber Förderdiagnostik. Feststellung des sonderpädagogischen Förderbedarfs im Förderschwerpunkt Lernen*. Hamburg: Persen.
- Helmke, A. (2009). Unterrichtsforschung. In K.-H. Arnold, U. Sandfuchs & J. Wiechmann (Hrsg.), *Handbuch Unterricht* (2. Aufl., S. 44–50). Bad Heilbrunn: Klinkhardt.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie, Band 3 (Psychologie der Schule und des Unterrichts)* (S. 71–176). Göttingen: Hogrefe.
- Henke, T., Bosse, S., Lambrecht, J., Jäntsche, C., Jaeuthe, J. & Spörer, N. (2017). Mittendrin oder nur dabei? Zum Zusammenhang zwischen sonderpädagogischem Förderbedarf und sozialer Partizipation von Grundschülerinnen und Grundschülern. *Zeitschrift für Pädagogische Psychologie*, 31, 111–123. <https://doi.org/10.1024/1010-0652/a000196>
- Heydrich, J., Weinert, S., Nusser, L., Artelt, C. & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: Challenges and approaches within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online*, 5, 217–240.
- Hibel, J., Farkas, G. & Morgan, P. L. (2010). Who is placed into special education? *Sociology of Education*, 83, 312–332. <https://doi.org/10.1177/0038040710383518>
- Hinz, A. (2002). Von der Integration zur Inklusion - terminologisches Spiel oder konzeptionelle Weiterentwicklung? *Zeitschrift für Heilpädagogik*, 53, 354–361.
- Hinz, A. (2004). From segregation to inclusion in Germany. In V. Heung & M. Ainscow (Hrsg.), *Inclusive education: A framework for reform* (S. 135–145). Hong Kong: Institute of Education.
- Hinz, A. (2009). Inklusive Pädagogik in der Schule - veränderter Orientierungsrahmen für die schulische Sonderpädagogik!? Oder doch deren Ende?? *Zeitschrift für Heilpädagogik*, 5, 171–179.
- Hinz, A. (2010). Towards inclusive education in Germany – structures, practical and theoretical development of joint education. In G. Bunch & A. Valeo (Hrsg.), *Inclusive education: Emergent solutions* (S. 40–73). Toronto: Inclusion Press.
- Ho, D. E., Imai, K., King, G. & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15, 199–236.
<https://doi.org/10.1093/pan/mpi013>
- Hocutt, A. M. (1996). Effectiveness of special education: Is placement the critical factor? *The Future of Children*, 6, 77–102. <https://doi.org/10.2307/1602495>
- Hofer, M. (2014). Persönlichkeitsentwicklung als schulisches Erziehungsziel jenseits kognitiver Kompetenzen. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 46, 55–66.
- Holland, P. W. & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Hrsg.), *Test validity* (S. 129–145). Hillsdale, NJ: Erlbaum.
- Hopf, W. & Kronauer, M. (2016). Welche Inklusion? In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 14–26). Weinheim: Beltz Juventa.
- Hornstra, L., Denessen, E., Bakker, J., Bergh, L. van den & Voeten, M. (2010). Teacher attitudes toward dyslexia: Effects on teacher expectations and the academic achievement of students with dyslexia. *Journal of Learning Disabilities*, 43, 515–529.
<https://doi.org/10.1177/0022219409355479>
- Horster, D. (2015). Bildungsgerechtigkeit aus sozialphilosophischer Sicht. In V. Maniatis, B. Hermstein, N. Berkemeyer & W. Bos (Hrsg.), *Zur Gerechtigkeit von Schule: Theorien, Konzepte, Analysen* (S. 42–50). Münster: Waxmann.
- Huber, C. (2009). Soziale Ausgrenzung in der Integration von Schülern mit sonderpädagogischem Förderbedarf. Zusammenhang von Persönlichkeit, Gruppenheterogenität und sozialer Ausgrenzung. *Empirische Pädagogik*, 23, 170–190.

- Huynh, H., Meyer, J. P. & Gallant, D. J. (2004). Comparability of student performance between regular and oral administrations for a high-stakes mathematics test. *Applied Measurement in Education*, 17, 39–57.
- Jackson, M., Erikson, R., Goldthorpe, J. H. & Yaish, M. (2007). Primary and secondary effects in class differentials in educational attainment: The transition to A-level courses in England and Wales. *Acta Sociologica*, 50, 211–229.
- Jahoda, M., Lazarsfeld, P. F. & Zeisel, H. (1975). *Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit* (25. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Jerusalem, M. (1997). Schulklasseneffekte. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie, Band 3 (Psychologie der Schule und des Unterrichts)* (S. 253–278). Göttingen: Hogrefe.
- Jerusalem, M., Drössler, S., Kleine, D., Klein-Heßling, J., Mittag, W. & Röder, B. (2009). *Förderung von Selbstwirksamkeit und Selbstbestimmung. Skalen zur Erfassung von Lehrer und Schülermerkmalen im Unterricht*. Berlin: Humboldt-Universität zu Berlin.
- Joncas, M. & Foy, P. (2012). Sample design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Hrsg.), *Methods and procedures in TIMSS and PIRLS 2011* (S. 1–21). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Justice, L. M., Logan, J. A. R., Lin, T.-J. & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, 25, 1722–1729. <https://doi.org/10.1177/0956797614538978>
- Kany, W. & Schöler, H. (2009). *Diagnostik schulischer Lern- und Leistungsschwierigkeiten. Ein Leitfaden*. Stuttgart: Kohlhammer.
- Kato, K., Moen, R. E. & Thurlow, M. L. (2009). Differentials of a state reading assessment: Item functioning, distractor functioning, and omission frequency for disability categories. *Educational Measurement: Issues and Practice*, 28(2), 28–40. <https://doi.org/10.1111/j.1745-3992.2009.00145.x>
- Kaufman, A. S. & Kaufman, N. L. (Hrsg.). (2009). *Kaufman Assessment Battery for Children (K-ABC)* (8. Aufl.). Göttingen: Hogrefe Verlag.
- Kemper, T. & Weishaupt, H. (2011). Zur Bildungsbeteiligung ausländischer Schüler an Förderschulen - unter besonderer Berücksichtigung der spezifischen Staatsangehörigkeit. *Zeitschrift für Heilpädagogik*, 62, 419–431.
- Kettler, R. J. & Elliott, S. N. (2010). Assessment accommodations for children with special needs. In B. McGaw, E. Baker & P. Peterson (Hrsg.), *International encyclopedia of education* (3. Aufl.). Oxford, UK: Elsevier.
- Klauer, K. J. & Lauth, G. W. (1997). Lernbehinderungen und Leistungsschwierigkeiten bei Schülern. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie, Band 3 (Psychologie der Schule und des Unterrichts)* (S. 701–738). Göttingen: Hogrefe.
- Klein, G. (2001). Sozialer Hintergrund und Schullaufbahn von Lernbehinderten/ Förderschülern 1969 und 1997. *Zeitschrift für Heilpädagogik*, 52, 51–61.
- Klieme, E. & Vieluf, S. (2013). Schulische Bildung im internationalen Vergleich. Ein Rahmenmodell für Kontextanalysen in PISA. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 - Impulse für die Schul- und Unterrichtsforschung* (S. 229–246). Weinheim: Beltz.
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1994). *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluß vom 06.05.1994*. Zugriff am 31.05.2013 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopae94.pdf>
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1998). *Empfehlungen zum Förderschwerpunkt Sprache. Beschluss vom 26.06.1998*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sprache.pdf>
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1999). *Empfehlungen zum Förderschwerpunkt Lernen. Beschluss vom 01.10.1999*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopale.pdf>

- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2000). *Empfehlungen zum Förderschwerpunkt emotionale und soziale Entwicklung. Beschluss vom 10.03.2000*. Zugriff am 30.05.2013 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/emotsozentw.pdf>
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010). *Pädagogische und rechtliche Aspekte der Umsetzung des Übereinkommens der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (Behindertenrechtskonvention - VN-BRK) in der schulischen Bildung. Beschluss vom 18.11.2010*. Zugriff am 03.08.2017 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_11_18-Behindertenrechtskonvention.pdf
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2011). *Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen. Beschluss vom 20.10.2011*. Zugriff am 02.09.2017 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Definitionenkatalog zur Schulstatistik 2015 (2)*. Zugriff am 18.09.2017 unter http://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Defkat2015_2.pdf
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Sonderpädagogische Förderung in Schulen 2005 bis 2014*. Zugriff am 03.08.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf
- Knigge, M. (2017). Wie ein negatives gesellschaftliches Ansehen der besuchten Schulform das Lernen erschweren kann. *Sonderpädagogische Förderung in Brandenburg, 1*, 26–33.
- Köbsell, S. (2012). Integration/Inklusion aus Sicht der Disability Studies: Aspekte aus der internationalen und der deutschen Diskussion. In K. Rathgeb (Hrsg.), *Disability Studies. Kritische Perspektiven für die Arbeit am Sozialen* (S. 39–54). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kocaj, A., Kuhl, P. & Stanat, P. (2016, März). *Auswirkungen des Gemeinsamen Unterrichts auf Kinder ohne sonderpädagogischen Förderbedarf*. Konferenzbeitrag gehalten auf der 4. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Berlin.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Köller, O. (2010). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (3. Aufl., S. 529–548). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O. & Baumert, J. (2001). Leistungsgruppierungen in der Sekundarstufe I und ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie, 15*, 99–110. <https://doi.org/10.1024/1010-0652.15.2.99>
- Köller, O. & Baumert, J. (2012). Schulische Leistungen und ihre Messung. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (7. Aufl., S. 645–661). Weinheim: Beltz.
- Köller, O., Schütte, K., Zimmermann, F., Retelsdorf, J. & Leucht, M. (2013). Starke Klasse, hohe Leistungen? Die Rolle der Leistungsstärke der Klasse für die individuellen Mathematik- und Leseleistungen in der Sekundarstufe I. *Psychologie in Erziehung und Unterricht, 60*, 184–197. <https://doi.org/10.2378/peu2013.art15d>
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für pädagogische Psychologie, 20*, 27–39. <https://doi.org/10.1024/1010-0652.20.1.27>
- Kornmann, R. (2006). Die Überrepräsentation ausländischer Kinder und Jugendlicher in Sonderschulen mit dem Schwerpunkt Lernen. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (S. 71–85). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kottmann, B. (2006). *Selektion in die Sonderschule: Das Verfahren zur Feststellung von sonderpädagogischem Förderbedarf als Gegenstand empirischer Forschung*. Bad Heilbrunn: Klinkhardt.
- Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M. & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers.

- Journal of Educational Psychology*, 100, 716–725. <https://doi.org/10.1037/0022-0663.100.3.716>
- Kretschmann, R. (2006). Diagnostik bei Lernbehinderungen. In U. Petermann & F. Petermann (Hrsg.), *Diagnostik sonderpädagogischen Förderbedarfs* (S. 139–162). Göttingen: Hogrefe Verlag.
- Kronig, W. (2001). Probleme der Selektion in den Grundschuljahren. *Schweizerische Zeitschrift für Bildungswissenschaften*, 23, 357–364.
- Kronig, W. (2003). Das Konstrukt des leistungsschwachen Immigrantenkindes. *Zeitschrift für Erziehungswissenschaft*, 6, 126–141. <https://doi.org/10.1007/s11618-003-0008-3>
- Lane, S. & Leventhal, B. (2015). Psychometric challenges in assessing English language learners and students with disabilities. *Review of Research in Education*, 39, 165–214. <https://doi.org/10.3102/0091732X14556073>
- Latham & Watkins. (2014). *Schulgesetze - Synopsen*. Zugriff am 24.03.2017 unter <http://www.gemeinsamleben-gemeinsamlernen.de/node/6>
- Lauth, G. W., Grünke, M. & Brunstein, J. C. (Hrsg.). (2014). *Interventionen bei Lernstörungen. Förderung, Training und Therapie in der Praxis* (2. Aufl.). Göttingen: Hogrefe.
- Legault, L., Green-Demers, I. & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, 98, 567–582. <https://doi.org/10.1037/0022-0663.98.3.567>
- Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 64, 123–153. <https://doi.org/10.1007/s11577-012-0158-5>
- Lehmann, R. & Hoffmann, E. (Hrsg.). (2009). *BELLA: Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf „Lernen“*. Münster: Waxmann.
- Lentz, A. & Radtke, F.-O. (1994). Bildungsghettos - Institutionalisierte Diskriminierung von Migrantenkindern in der Grundschule. *Unterrichtswissenschaft*, 22, 182–191.
- Lindsay, G. (2007). Educational psychology and the effectiveness of inclusive education/mainstreaming. *British Journal of Educational Psychology*, 77, 1–24. <https://doi.org/10.1348/000709906X156881>
- Lindsay, G., Pather, S. & Strand, S. (2006). *Special educational needs and ethnicity: Issues of over- and under-representation*. Annesley, Nottingham: Department for Education and Skills.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, 30, 263–285. <https://doi.org/10.1016/j.cedpsych.2004.10.002>
- Lüdtke, O., Robitzsch, A. & Köller, O. (2002). Statistische Artefakte bei Kontexteffekten in der pädagogisch-psychologischen Forschung. *Zeitschrift für Pädagogische Psychologie*, 16, 217–231. <https://doi.org/10.1024/1010-0652.16.34.217>
- Luhmann, N. & Schorr, K. E. (1988). *Reflexionsprobleme im Erziehungssystem* (4. Aufl.). Frankfurt am Main: Suhrkamp Verlag.
- Maaz, K., Baumert, J. & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (S. 11–46). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maaz, K., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2, 99–106. <https://doi.org/10.1111/j.1750-8606.2008.00048.x>
- Mace, R. L. (1991). *Definitions: Accessible, adaptable, and universal design (Fact Sheet)*. Raleigh, NC: Center for Universal Design, NC State University.
- Malecki, A. (2013). Sonderpädagogische Förderung in Deutschland – eine Analyse der Datenlage in der Schulstatistik. *WISTA – Wirtschaft und Statistik*, 5, 356–365.
- Malecki, A. (2014). Sonderpädagogischer Förderbedarf – eine differenzierte Analyse. *WISTA – Wirtschaft und Statistik*, 10, 591–601.

- Mand, J. (2003). Gemeinsamer Unterricht oder Sonderschule? Diagnostische Arbeit unter Alltagsbedingungen. In G. Feuser (Hrsg.), *Integration heute. Perspektiven ihrer Weiterentwicklung in Theorie und Praxis*. (S. 143–152). Frankfurt, Main: Lang.
- Markussen, E. (2004). Special education: Does it help? A study of special education in Norwegian upper secondary schools. *European Journal of Special Needs Education*, 19, 33–48.
<https://doi.org/10.1080/0885625032000167133>
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23, 129–149.
<https://doi.org/10.3102/00028312023001129>
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280–295. <https://doi.org/10.1037/0022-0663.79.3.280>
- Marsh, H. W. (2005). *Big fish little pond effect on academic self-concept : Cross-cultural and cross-disciplinary generalizability*. Konferenzbeitrag gehalten auf der Australian Association for Research in Education 2005, Parramatta, N.S.W. Konferenzbeitrag abgerufen von <https://www.aare.edu.au/data/publications/2005/mar05389.pdf>
- Marsh, H. W., Kong, C. K. & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology*, 78, 337–349.
<https://doi.org/10.1037/0022-3514.78.2.337>
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J. & Craven, R. G. (2008). The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20, 319–350.
<https://doi.org/10.1007/s10648-008-9075-6>
- Marsh, H. W., Tracey, D. K. & Craven, R. G. (2006). Multidimensional self-concept structure for preadolescents with mild intellectual disabilities: A hybrid multigroup–MIMC approach to factorial invariance and latent mean differences. *Educational and Psychological Measurement*, 66, 795–818. <https://doi.org/10.1177/0013164405285910>
- Marx, P. & Lenhard, W. (2011). Diagnostische Merkmale von Screening-Verfahren zur Früherkennung möglicher Probleme beim Schriftspracherwerb. In M. Hasselhorn & W. Schneider (Hrsg.), *Frühprognose schulischer Kompetenzen* (S. 68–84). Göttingen: Hogrefe.
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41, 121–131. <https://doi.org/10.1026/0049-8637.41.3.121>
- McLeskey, J. & Waldron, N. L. (2011). Educational programs for elementary students with learning disabilities: Can they be both effective and inclusive? *Learning Disabilities Research & Practice*, 26, 48–57. <https://doi.org/10.1111/j.1540-5826.2010.00324.x>
- Meece, J. L., Anderman, E. M. & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, 57, 487–503.
<https://doi.org/10.1146/annurev.psych.56.091103.070258>
- Meyer, J. W. (1977). The effects of education as an institution. *American Journal of Sociology*, 83, 55–77. <https://doi.org/10.1086/226506>
- Mitchell, D. (2013). *What really works in special and inclusive education: Using evidence-based teaching strategies* (2. Aufl.). London: Routledge.
- Möller, J. & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, 120, 544–560. <https://doi.org/10.1037/a0032459>
- Möller, J., Streblow, L. & Pohlmann, B. (2009). Achievement and self-concept of students with learning disabilities. *Social Psychology of Education*, 12, 113–122.
<https://doi.org/10.1007/s11218-008-9065-z>
- Morgan, P. L., Farkas, G. & Hibell, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, 31, 187–198.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H. & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher*, 44, 278–292.
<https://doi.org/10.3102/0013189X15591157>

- Morgan, P. L., Farkas, G. & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities*, 44, 472–488. <https://doi.org/10.1177/0022219411414010>
- Moser, V. (2012). Braucht die Inklusionspädagogik einen Behinderungsbegriff? *Zeitschrift für Inklusion*, 3. Zugriff am 27.10.2016 unter <https://www.inklusion-online.net/index.php/inklusion-online/article/view/40>
- Müller, K., Sälzer, C., Mang, J., Gebhardt, M. & Prenzel, M. (2013, März). *Aufgabenbezogenes Aufmerksamkeitsverhalten von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf*. Konferenzbeitrag gehalten auf der 1. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Kiel.
- Müller, K., Sälzer, C., Mang, J., Heine, J.-H., Prenzel, M. & Gebhardt, M. (2014, März). *Kompetenzen von Schülerinnen und Schüler mit besonderem Förderbedarf. Ergebnisse aus dem PISA 2012 Förderschul-Oversample zum Nutzen von differentiellen Maßnahmen der Testadaptation*. Konferenzbeitrag gehalten auf der 2. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Frankfurt a.M.
- Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T. & Trautwein, U. (2011). Who took the “×” out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, 22, 1058–1066. <https://doi.org/10.1177/0956797611415540>
- Naraian, S. (2011). Seeking transparency: The production of an inclusive classroom community. *International Journal of Inclusive Education*, 15, 955–973. <https://doi.org/10.1080/13603110903477397>
- National Center for Educational Statistics. (2010). *NAEP Testing and Reporting on Students with Disabilities and English Language Learners. Policy Statement*. Washington, DC: National Center for Education Statistics.
- Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus*. Münster: Waxmann.
- Nusser, L., Carstensen, C. H. & Artelt, C. (2015). Befragung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen: Ergebnisse zur Messinvarianz. *Empirische Sonderpädagogik*, 7, 99–116.
- Nusser, L. & Wolter, I. (2016). There's plenty more fish in the sea : Das akademische Selbstkonzept von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen in integrativen und segregierten Schulsettings. *Empirische Pädagogik*, 30, 130–143.
- OECD = Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results: What makes a school successful (Volume IV): Resources, policies and practices*. Paris: OECD Publishing.
- OECD = Organisation for Economic Co-operation and Development. (2014). *PISA 2012 technical report*. Paris: OECD Publishing.
- Olson, J. F. & Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress* (Research and Development Report No. NCES 97-482). U.S. Department of Education, Office of Educational Research and Improvement.
- Pallas, A. M., Entwisle, D. R., Alexander, K. L. & Stuka, M. F. (1994). Ability-group effects: Instructional, social, or institutional? *Sociology of Education*, 67, 27–46. <https://doi.org/10.2307/2112748>
- Pant, H. A. (2015, Oktober). *Wie kann Inklusion gelingen?* Vortrag gehalten auf dem Bildungspolitischen Forum des Leibniz-Forschungsverbunds Bildungspotenziale „Akzeptanz und Umsetzbarkeit von Bildungsreformen: Wie wir Bildungspotenziale ausschöpfen können“, Berlin.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Patrick, H., Ryan, A. M. & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, 99, 83–98. <https://doi.org/10.1037/0022-0663.99.1.83>

- Pearl, R. & Bay, M. (1999). Psychosocial correlates of learning disabilities. In V. L. Schwann & D. H. Saklofske (Hrsg.), *Handbook of psychosocial characteristics of exceptional children* (S. 443–470). New York: Kluwer/Plenum.
- Peetsma, T., Vergeer, M., Roeleveld, J. & Karsten, S. (2001). Inclusion in education: Comparing pupils' development in special and regular education. *Educational Review*, 53, 125–135. <https://doi.org/10.1080/00131910125044>
- Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, 41, 359–376. <https://doi.org/10.1111/j.1464-0597.1992.tb00712.x>
- Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91–105. https://doi.org/10.1207/S15326985EP3702_4
- Penfield, R. D. & Camilli, G. (2007). Differential item functioning and item bias. In C.R. Rao and S. Sinharay (Hrsg.), *Handbook of Statistics* (Bd. Volume 26, S. 125–167). Amsterdam: Elsevier/North Holland.
- Petermann, U., Menzel, S. & Eckert, F. (2003). Verhaltensbeurteilungsbogen Schule (VBS-L). *Zeitschrift für Heilpädagogik*, 54, 99–104.
- Petermann, U. & Petermann, F. (Hrsg.). (2006). *Diagnostik sonderpädagogischen Förderbedarfs*. Göttingen: Hogrefe Verlag.
- Pfahl, L. (2011). *Techniken der Behinderung: Der deutsche Lernbehinderungsdiskurs, die Sonderschule und ihre Auswirkungen auf Bildungsbiografien* (1., Aufl.). Bielefeld: Transcript.
- Pfahl, L. (2012). Bildung, Behinderung und Agency. Eine wissenssoziologische Untersuchung der Folgen schulischer Segregation und Inklusion. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung* (S. 415–436). Wiesbaden: Springer VS.
- Pfahl, L. & Powell, J. J. W. (2011). Legitimizing school segregation. The special education profession and the discourse of learning disability in Germany. *Disability & Society*, 26, 449–462. <https://doi.org/10.1080/09687599.2011.567796>
- Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.
- Pitoniak, M. J. & Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, 71, 53–104.
- Pohl, S., Südkamp, A., Hardt, K., Carstensen, C. H. & Weinert, S. (2016). Testing students with special educational needs in large-scale assessments - psychometric properties of test scores and associations with test taking behavior. *Frontiers in Psychology*, 7(154), 1–14. <https://doi.org/10.3389/fpsyg.2016.00154>
- Poscher, R., Rux, J. & Langer, T. (2008). *Von der Integration zur Inklusion. Das Recht auf Bildung aus der Behindertenrechtskonvention der Vereinten Nationen und seine innerstaatliche Umsetzung*. Baden-Baden: Nomos.
- Powell, J. J. W. (2004). *Barriers to inclusion: The institutionalization of special education in the United States and Germany* (Dissertation, Freie Universität Berlin, Berlin, Deutschland). Abgerufen von <http://orbilu.uni.lu/handle/10993/5477>
- Powell, J. J. W. (2009). To segregate or to separate? Special education expansion and divergence in the United States and Germany. *Comparative Education Review*, 53, 161–187. <https://doi.org/10.1086/597816>
- Powell, J. J. W. & Pfahl, L. (2012). Sonderpädagogische Fördersysteme. In U. Bauer, U. H. Bittlingmayer & A. Scherr (Hrsg.), *Handbuch Bildungs- und Erziehungssoziologie* (S. 721–739). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Powell, J. J. W. & Wagner, S. J. (2002). Zur Entwicklung der Überrepräsentanz von Migrantenjugendlichen an Sonderschulen in der BRD seit 1991. *Gemeinsam Leben*, 10(2), 66–71.
- Prenzel, A. (2013). *Inklusive Bildung in der Primarstufe. Eine wissenschaftliche Expertise für den Grundschulverband*. Frankfurt: Grundschulverband.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

- Preuss-Lausitz, U. (2001). Gemeinsamer Unterricht Behinderter und Nichtbehinderter. *Zeitschrift für Erziehungswissenschaft*, 4, 209–224. <https://doi.org/10.1007/s11618-001-0024-0>
- Preuss-Lausitz, U. (2009). Rezension zu: Lehmann, R./Hoffmann, E. (Hrsg.): Bella. Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf Lernen. *Zeitschrift für Heilpädagogik*, 60(11), 46–47.
- Preuss-Lausitz, U. (2013). Inklusive Schul- und Unterrichtsentwicklung. Ein Beitrag zum Abbau von Ungleichheit. In E. Jürgens & S. Miller (Hrsg.), *Ungleichheit in der Gesellschaft und Ungleichheit in der Schule: eine interdisziplinäre Sicht auf Inklusions- und Exklusionsprozesse* (S. 171–187). Weinheim: Beltz Juventa.
- Preuss-Lausitz, U. (2014). *Wissenschaftliche Begleitungen der Wege zur inklusiven Schulentwicklung in den Bundesländern. Versuch einer Übersicht*. Zugriff am 20.10.2015 unter https://www.ewi.tu-berlin.de/fileadmin/i49/dokumente/Preuss-Lausitz/Wiss._Begleitung_Inklusion.pdf
- Raudenbush, S. W. & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- Rauer, W. & Schuck, K. D. (2007). Hamburger Grundschulen und Grundschulklassen mit einer formellen Integrationsorganisation. In W. Bos, C. Gröhlich & M. Pietsch (Hrsg.), *KESS 4 – Lehr- und Lernbedingungen in Hamburger Grundschulen*. Münster: Waxmann.
- Ready, D. D. & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48, 335–360. <https://doi.org/10.3102/0002831210374874>
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim: Beltz.
- Renick, M. J. & Harter, S. (1989). Impact of social comparisons on the developing self-perceptions of learning disabled students. *Journal of Educational Psychology*, 81, 631–638. <https://doi.org/10.1037/0022-0663.81.4.631>
- Rheinberg, F. & Enstrup, B. (1977). Self-concept of intelligence with pupils in special and normal schools: A group effect. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9, 171–180.
- Richter, D., Engelbert, M., Böhme, K., Haag, N., Hannighofer, J., Reimers, H., ... Stanat, P. (2012). Anlage und Durchführung des Ländervergleichs. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011* (S. 85–102). Münster: Waxmann.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Münster: Waxmann.
- Ricking, H. (2005). Der „Overlap“ von Lern- und Verhaltensstörungen. *Sonderpädagogik*, 35, 235–248.
- Rist, R. C. (2000). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education (HER classic reprint). *Harvard Educational Review*, 70, 257–302.
- Rix, J. & Sheehy, K. (2014). Nothing special: The everyday pedagogy of teaching. In L. Florian (Hrsg.), *The SAGE Handbook of Special Education: Two Volume Set* (2. Aufl., Bde. 1–2, S. 459–474). London: SAGE.
- Rogers, C. M., Christian, E. M. & Thurlow, M. L. (2012). *A summary of the research on the effects of test accommodations: 2009-2010 (Technical Report 65)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Rogers, D. P. & Thiery, I. M. (2003, November). *Does an inclusive setting affect reading comprehension in students with learning disabilities?* Konferenzbeitrag gehalten auf dem Annual Meeting of the Mid-South Educational Research Association, Biloxi, MS. Konferenzbeitrag abgerufen von <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED482684>

- Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39, 33–38. <https://doi.org/10.1080/00031305.1985.10479383>
- Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, 3, 16–20. <https://doi.org/10.1007/BF02322211>
- Rossmann, P., Klicpera Gasteiger, B., Gebhardt, M., Roloff, C. & Weindl, A. (2011). Zum Selbstkonzept von SchülerInnen mit Sonderpädagogischem Förderbedarf in Sonderschulen und Integrationsklassen: Ein empirisch fundierter Diskussionsbeitrag. In R. Mikula & H. Kittl-Satran (Hrsg.), *Dimensionen der Erziehungs- und Bildungswissenschaft* (S. 107–120). Graz: Leykam.
- Ruijs, N. M. & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4, 67–79. <https://doi.org/10.1016/j.edurev.2009.02.002>
- Ruijs, N., Peetsma, T. & van der Veen, I. (2010). The presence of several students with special educational needs in inclusive education and the functioning of students with special educational needs. *Educational Review*, 62, 1–37. <https://doi.org/10.1080/00131910903469551>
- Rumberger, R. W. & Palardy, G. J. (2005). Test scores, dropout rates, and transfer rates as alternative indicators of high school performance. *American Educational Research Journal*, 42, 3–42. <https://doi.org/10.3102/00028312042001003>
- Russell, S. L. (2012). *Individual- and classroom-level social support and classroom behavior in middle school* (Dissertation, University of Maryland, College Park, USA). Abgerufen von <http://drum.lib.umd.edu/handle/1903/13204>
- Rutter, M. (1983). School effects on pupil progress: Research findings and policy implications. *Child Development*, 54, 1–29. <https://doi.org/10.2307/1129857>
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 1, 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Wiesbaden: Springer VS.
- Sauter, S. (2013). Bildung für alle – Schule für alle? - Ausblicke auf ein schulpädagogisches Spannungsfeld im Kontext von Steuerungslogik und der aktuellen Debatte um Inklusion. *Zeitschrift für Inklusion*, 1. Zugriff am 03.07.2017 unter <https://www.inklusion-online.net/index.php/inklusion-online/article/view/27>
- Schafer, J. L. & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods*, 13, 279–313. <https://doi.org/10.1037/a0014268>
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung: Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Münster: Waxmann.
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1, 61–80. <https://doi.org/10.1080/0924345900010106>
- Scheerens, J. (2004). *Review of school and instructional effectiveness research* (Paper commissioned for the EFA Global Monitoring Report 2005, The Quality Imperative). Zugriff am 10.11.2016 unter <http://unesdoc.unesco.org/images/0014/001466/146695e.pdf>
- Scheerens, J. (2008). *Review of research on school and instructional effectiveness*. Enschede: University of Twente, Department of Educational Organization and Management.
- Schofield, J. W. (2010). International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. *Teachers College Record*, 112, 1492–1528.
- Schor, B. J. (2003). Die Differenziertheit des deutschen Bildungswesens. Chance oder Hemmnis. *Zeitschrift für Heilpädagogik*, 54, 370–376.
- Schründer-Lenzen, A. & Krückels, C. (2015). Die Grundschule im Land Brandenburg. Frühe Reformansätze und aktuelle Strukturen. In N. Spörer, A. Schründer-Lenzen, M. Vock & ge-

- meinsam mit dem Landesinstitut für Schule und Medien Berlin Brandenburg (LISUM) (Hrsg.), *Inklusives Lernen und Lehren im Land Brandenburg. Abschlussbericht zur Begleitforschung des Pilotprojekts „Inklusive Grundschule“* (S. 7–24). Ludwigsfelde, Potsdam: Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM), Universität Potsdam.
- Schumann, B. (2007). „*Ich schäme mich ja so!*“. *Die Sonderschule für Lernbehinderte als „Schonraumfalle“*. Bad Heilbrunn: Klinkhardt.
- Schümer, G. (2004). Zur doppelten Benachteiligung von Schülern aus unterprivilegierten Gesellschaftsschichten im deutschen Schulwesen. In G. Schümer, K.-J. Tillmann & M. Weiss (Hrsg.), *Die Institution Schule und die Lebenswelt der Schüler. Vertiefende Analysen der PISA-2000-Daten zum Kontext von Schülerleistungen* (1. Aufl, S. 73–114). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schütz, G., West, M. R. & Wößmann, L. (2007). *School accountability, autonomy, choice, and the equity of student achievement* (OECD Education Working Papers). Paris: Organisation for Economic Co-operation and Development.
- Schwinger, M., Grunsel, C., Yotyodying, S., Baumanns, R., Gorges, J. & Serke, B. (2014, März). *Hat inklusive vs. exklusive Beschulung einen Einfluss auf Motivation, Selbstwert und Wohlbefinden von Kindern mit sonderpädagogischem Förderbedarf? Erste Befunde der Bielefelder Längsschnittstudie BiLieF*. Konferenzbeitrag gehalten auf der 2. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Frankfurt a.M.
- Schwinger, M., Wild, E., Lütje-Klose, B., Grunsel, C., Stranghöner, D., Yotyodying, S., ... Stelling, S. (2015). Wie können motivationale und affektive Merkmale bei Kindern mit sonderpädagogischem Förderbedarf valide erfasst werden? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 273–300). Wiesbaden: Springer VS.
- Scruggs, T. E., Mastropieri, M. A. & Marshak, L. (2012). Peer-mediated instruction in inclusive secondary social studies learning: direct and indirect learning effects. *Learning Disabilities Research & Practice*, 27, 12–20. <https://doi.org/10.1111/j.1540-5826.2011.00346.x>
- Scruggs, T. E., Mastropieri, M. A. & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional Children*, 73, 392–416. <https://doi.org/10.1177/001440290707300401>
- Senatsverwaltung für Bildung, Jugend und Wissenschaft. (2012). *Leitfaden zur Feststellung sonderpädagogischen Förderbedarfs an Berliner Schulen*. Berlin: Senatsverwaltung für Bildung, Jugend und Wissenschaft.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407–441. <https://doi.org/10.2307/1170010>
- Shepard, L. A. (1993). Evaluating test validity. *Review of Research in Education*, 19, 405–450.
- Sireci, S. G. (2004). *Validity issues in accommodating NAEP reading tests*. Washington, DC: National Assessment Governing Board.
- Sireci, S. G. (2009). No more excuses: New research on assessing students with disabilities. *Journal of Applied Testing Technology*, 10(2), 1–18.
- Sireci, S. G., Scarpati, S. E. & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75, 457–490. <https://doi.org/10.3102/00346543075004457>
- Skinner, E. A. & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571–581. <https://doi.org/10.1037/0022-0663.85.4.571>
- Smith, E. & Douglas, G. (2014). Special educational needs, disability and school accountability: an international perspective. *International Journal of Inclusive Education*, 18, 443–458. <https://doi.org/10.1080/13603116.2013.788222>
- Solga, H. (2005). Meritokratie - die moderne Legitimation ungleicher Bildungschancen. In P. A. Berger & H. Kahlert (Hrsg.), *Institutionalisierte Ungleichheiten: Wie das Bildungswesen Chancen blockiert* (S. 19–38). Weinheim: Juventa.

- Song, J., Bong, M., Lee, K. & Kim, S. (2015). Longitudinal investigation into the role of perceived social support in adolescents' academic motivation and achievement. *Journal of Educational Psychology*, 107, 821–841. <https://doi.org/10.1037/edu0000016>
- Sozialgesetzbuch (SGB) Neuntes Buch (IX) - Rehabilitation und Teilhabe behinderter Menschen in der Fassung des Bundesteilhabegesetzes vom 23. Dezember 2016 (BGBl. I S. 2334), zuletzt geändert durch Artikel 27 des Gesetzes vom 17. Juli 2017.
- Spörer, N., Maaz, K., Vock, M., Schröder-Lenzen, A., Luka, T., Bosse, S., ... Jäntsche, C. (2015). Lernen in der inklusiven Grundschule: Zusammenhänge zwischen fachlichen Kompetenzen, Sozialklima und Facetten des Selbstkonzepts. *Unterrichtswissenschaft*, 1, 22–35.
- Stäbler, F., Dumont, H., Becker, M. & Baumert, J. (2017). What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept. *Journal of Educational Psychology*, 109, 191–207. <https://doi.org/10.1037/edu0000135>
- Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Hrsg.). (2016). *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K., Richter, D., Weirich, S., Haag, N., Roppelt, A., Engelbert, M. & Reimers, H. (2014). *IQB Ländervergleich Primarstufe 2011 (IQB-LV 2011)*. Version: 2. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. http://doi.org/10.5159/IQB_LV_2011_v2
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.
- Stern, K.-H. (2005). Kinder und Jugendliche mit sonderpädagogischem Förderbedarf an der Regelschule. In S. Ellinger & M. Wittrock (Hrsg.), *Sonderpädagogik in der Regelschule. Konzepte-Forschung-Praxis*. Stuttgart: Kohlhammer.
- Stocké, V. (2007). Explaining educational decision and effects of families' social class position: An empirical test of the Breen–Goldthorpe model of educational attainment. *European Sociological Review*, 23, 505–519. <https://doi.org/10.1093/esr/jcm014>
- Story, M. F., Mueller, J. L. & Mace, R. L. (1998). *The universal design file: Designing for people of all ages and abilities. Revised edition*. Raleigh, NC: Center for Universal Design, NC State University.
- Stranghöner, D., Hollmann, J., Otterpohl, N., Wild, E., Lütje-Klose, B. & Schwinger, M. (2017). Inklusion versus Exklusion: Schulsetting und Lese-Rechtschreibentwicklung von Kindern mit Förderschwerpunkt Lernen. *Zeitschrift für Pädagogische Psychologie*, 31, 125–136. <https://doi.org/10.1024/1010-0652/a000202>
- Südkamp, A., Pohl, S., Hardt, K., Jordan, A.-K. & Duchhardt, C. (2015). Kompetenzmessung in den Bereichen Lesen und Mathematik bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 243–272). Wiesbaden: Springer VS.
- Südkamp, A., Pohl, S. & Weinert, S. (2015). Competence assessment of students with special educational needs - Identification of appropriate testing accommodations. *Frontline Learning Research*, 3(2), 1–25. <https://doi.org/10.14786/flr.v3i2.130>
- Szumski, G. & Karwowski, M. (2015). Emotional and social integration and the big-fish-little-pond effect among students with and without disabilities. *Learning and Individual Differences*, 43, 63–74. <https://doi.org/10.1016/j.lindif.2015.08.037>
- Tent, L., Witt, M., Zschoche-Lieberum, C. & Bürger, W. (1991). Über die pädagogische Wirksamkeit der Schule für Lernbehinderte. *Zeitschrift für Heilpädagogik*, 5, 289–320.
- Thrupp, M., Lauder, H. & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37, 483–504. [https://doi.org/10.1016/S0883-0355\(03\)00016-8](https://doi.org/10.1016/S0883-0355(03)00016-8)
- Thurlow, M. L. (2002). Positive educational results for all students: The promise of standards-based reform. *Remedial and Special Education*, 23, 195–202.

- Thurlow, M. L. (2010). Steps toward creating fully accessible reading assessments. *Applied Measurement in Education*, 23, 121–131. <https://doi.org/10.1080/08957341003673765>
- Thurlow, M. L., Lazarus, S. S., Albus, D. & Hodgson, J. (2010). *Computer-based testing: Practices and considerations (Synthesis Report 78)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M. L., Ysseldyke, J., Gutman, S. & Geenen, K. (1998). *An analysis of inclusion of students with disabilities in state standards documents (Technical Report 19)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Tiekstra, M., Hessels, M. G. P. & Minnaert, A. E. M. G. (2009). Learning capacity in adolescents with mild intellectual disabilities. *Psychological Reports*, 105, 804–814. <https://doi.org/10.2466/PRO.105.3.804-814>
- Tiffin-Richards, S. P. & Pant, H. A. (2017). Arguing validity in educational assessment. In D. Leutner, J. Fleischer, J. Grünkorn & E. Klieme (Hrsg.), *Competence assessment in education. Research, models and instruments* (S. 469–485). Cham, Schweiz: Springer.
- Tillmann, K.-J. (2016). Empirische Bildungsforschung in der Kritik – ein Überblick über Themen und Kontroversen. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker* (S. 5–22). Wiesbaden: Springer VS.
- Tillmann, K.-J. & Wischer, B. (2006). Heterogenität in der Schule. Forschungsstand und Konsequenzen. *Pädagogik*, 58(3), 44–48.
- Tindal, G. & Anderson, D. (2011). Validity evidence for making decisions about accommodated and modified large-scale tests. In S. N. Elliott, R. J. Kettler, P. A. Beddow & A. Kurz (Hrsg.), *Handbook of accessible achievement tests for all students* (S. 183–200). New York, NY: Springer.
- Tomlinson, S. (2012). The irresistible rise of the SEN industry. *Oxford Review of Education*, 38, 267–286. <https://doi.org/10.1080/03054985.2012.692055>
- Tomlinson, S. (2015). Is a sociology of special and inclusive education possible? *Educational Review*, 67, 273–281. <https://doi.org/10.1080/00131911.2015.1021764>
- Tracey, D. K. & Marsh, H. W. (2000). Self-concepts of primary students with mild intellectual disabilities: Issues of measurement and educational placement. In R. G. Craven & H. W. Marsh (Hrsg.), *Self-concept theory, research and practice: Advances for the new millennium* (S. 419–425). Sydney: SELF Research Centre, University of Western Sydney.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O. & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, 98, 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>
- Trautwein, U., Lüdtke, O., Marsh, H. W. & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology*, 101, 853–866. <https://doi.org/10.1037/a0016306>
- UN BRK (2008). Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention) vom 21. 12. 2008.
- UN Committee on the Rights of Persons with Disabilities (Hrsg.). (2015). *Concluding observations on the initial report of Germany*. Geneva: United Nations.
- UNESCO = The United Nations Educational, Scientific and Cultural Organization. (1994). The Salamanca Statement and Framework for action on special needs education. Beschluss auf der World Conference on Special Needs Education; Access and Quality, Salamanca, Spain: UNESCO.
- UNESCO = The United Nations Educational, Scientific and Cultural Organization. (2012). *International Standard Classification of Education, ISCED 2011*. Montreal: UNESCO Institute for Statistics.
- Urdan, T. & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, 44, 331–349. <https://doi.org/10.1016/j.jsp.2006.04.003>

- Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39, 111–133. https://doi.org/10.1207/s15326985ep3902_3
- Vannest, K. J. & Hagan-Burke, S. (2010). Teacher time use in special education. *Remedial and Special Education*, 31, 126–142. <https://doi.org/10.1177/0741932508327459>
- Vannest, K. J. & Parker, R. I. (2010). Measuring time: The stability of special education teacher time use. *The Journal of Special Education*, 44, 94–106. <https://doi.org/10.1177/0022466908329826>
- Vaughn, S., Elbaum, B. E. & Schumm, J. S. (1996). The effects of inclusion on the social functioning of students with learning disabilities. *Journal of Learning Disabilities*, 29, 598–608.
- Wagner, S. J. & Powell, J. J. W. (2003). Ethnisch-kulturelle Ungleichheit im deutschen Bildungssystem. Zur Überrepräsentanz von Migrantenjugendlichen an Sonderschulen. In G. Cloerkes (Hrsg.), *Wie man behindert wird: Texte zur Konstruktion einer sozialen Rolle und zur Lebenssituation betroffener Menschen* (S. 183–208). Heidelberg: Winter.
- Wang, M. C. & Baker, E. T. (1985). Mainstreaming programs: Design features and effects. *The Journal of Special Education*, 19, 503–521. <https://doi.org/10.1177/002246698501900412>
- Wang, M.-T. & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12–23. <https://doi.org/10.1016/j.learninstruc.2013.04.002>
- Wei, X., Blackorby, J. & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, 78, 89–106.
- Weishaupt, H. (2016). Inklusion als umfassende schulische Innovation. Streitbare Anmerkungen zu einer wichtigen Schulreform. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion*. 62. Beiheft der Zeitschrift für Pädagogik (S. 27–41). Weinheim: Beltz Juventa.
- Weishaupt, H. (2017). Die Förderschule als sonderpädagogischer Förderort. In M. Gercke, S. Opalinski & T. Thonagel (Hrsg.), *Inklusive Bildung und gesellschaftliche Exklusion* (S. 39–57). Wiesbaden: Springer VS.
- Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology*, 90, 202–209. <https://doi.org/10.1037/0022-0663.90.2.202>
- Wentzel, K. R., Battle, A., Russell, S. L. & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35, 193–202. <https://doi.org/10.1016/j.cedpsych.2010.03.002>
- Werning, R. (2010). Inklusion – Herausforderungen, Widersprüche und Perspektiven. *Lernchancen*, 78, 4–9.
- Werning, R. (2014). Stichwort: Schulische Inklusion. *Zeitschrift für Erziehungswissenschaft*, 17, 601–623. <https://doi.org/10.1007/s11618-014-0581-7>
- Werning, R. & Lichtblau, M. (2012). Sonderpädagogische Diagnostik. In R. Werning, R. Balgo, W. Palmowski & M. Sassenroth (Hrsg.), *Sonderpädagogik: Lernen, Verhalten, Sprache, Bewegung und Wahrnehmung* (2. Aufl., S. 229–260). München: Oldenbourg Wissenschaftsverlag.
- Werning, R. & Lütje-Klose, B. (2006). *Einführung in die Lernbehindertenpädagogik* (2. Aufl.). München: Reinhardt.
- Werning, R. & Reiser, H. (2008). Sonderpädagogische Förderung. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 505–539). Reinbek bei Hamburg: Rowohlt.
- Wheeler, L. & Nezlek, J. (1977). Sex differences in social participation. *Journal of Personality and Social Psychology*, 35, 742–754. <https://doi.org/10.1037/0022-3514.76.3.420>
- Wheeler, L. & Suls, J. (2005). Social comparison and self-evaluations of competence. In A. J. Elliot & C. S. Dweck (Hrsg.), *Handbook of competence and motivation* (S. 566–578). New York, NY, US: Guilford Publications.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wild, E., Schwinger, M., Lütje-Klose, B., Yotyodying, S., Gorges, J., Stranghöner, D., ... Kurnitzki, S. (2015). Schülerinnen und Schüler mit dem Förderschwerpunkt Lernen in inklusiven und ex-

- klusiven Förderarrangements: Erste Befunde des BiLieF-Projektes zu Leistung, sozialer Integration, Motivation und Wohlbefinden. *Unterrichtswissenschaft*, 43, 7–21.
- Wilkinson, I. A. (2002). Peer influences on learning: where are they? *International Journal of Educational Research*, 37, 395–401.
- Wocken, H. (1996). Sonderpädagogischer Förderbedarf als systemischer Begriff. *Sonderpädagogik*, 1, 34–38.
- Wocken, H. (2000). Leistung, Intelligenz und Soziallage von Schülern mit Lernbehinderungen. Vergleichende Untersuchungen an Förderschulen in Hamburg. *Zeitschrift für Heilpädagogik*, 51, 492–503.
- Wocken, H. (2005). *Andere Länder, andere Schüler? Vergleichende Untersuchung von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen*. Zugriff am 03.08.2017 unter <http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf>
- Wocken, H. (2009). Inklusion & Integration. Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Träumereien zu bewahren. In A.-D. Stein, I. Niedek & S. Krach (Hrsg.), *Integration und Inklusion auf dem Wege ins Gemeinwesen* (S. 204–234). Bad Heilbrunn: Klinkhardt.
- Wocken, H. & Gröhlich, C. (2009). Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen. In W. Bos, M. Bonsen & C. Gröhlich (Hrsg.), *KESS 7 - Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (S. 133–142). Münster: Waxmann.
- World Health Organization. (2017). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. 10. Revision*. Geneva: World Health Organization.
- Wouters, S., Fraine, B., Colpin, H., Damme, J. & Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: A dynamic test of the big-fish-little-pond effect. *Journal of Educational Psychology*, 104, 793–805. <https://doi.org/10.1037/a0027732>
- Wrase, M. (2015). Die Implementation des Rechts auf inklusive Schulbildung nach der UN-Behindertenrechtskonvention und ihre Evaluation aus rechtlicher Perspektive. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 41–74). Wiesbaden: Springer VS.
- Wrase, M. (2016). *Auflösung der Förderschulen. Die UN-Behindertenkonvention verlangt die Inklusion von Kindern mit Behinderung an Regelschulen* (WZBrief Bildung No. 33). Berlin: Wissenschaftszentrum Berlin für Sozialforschung.
- Wygotski, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, Mass.: Harvard University Press.
- Zigmond, N. (2003). Where should students with disabilities receive special education services? Is one place better than another? *The Journal of Special Education*, 37, 193–199. <https://doi.org/10.1177/00224669030370030901>

Danksagung

Diese Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut zur Qualitätsentwicklung im Bildungswesen (IQB). Ich habe sehr vom kognitiv anregenden und netten Arbeitsumfeld profitiert und bedanke mich herzlich bei allen Menschen, die an dieser Arbeit beteiligt waren.

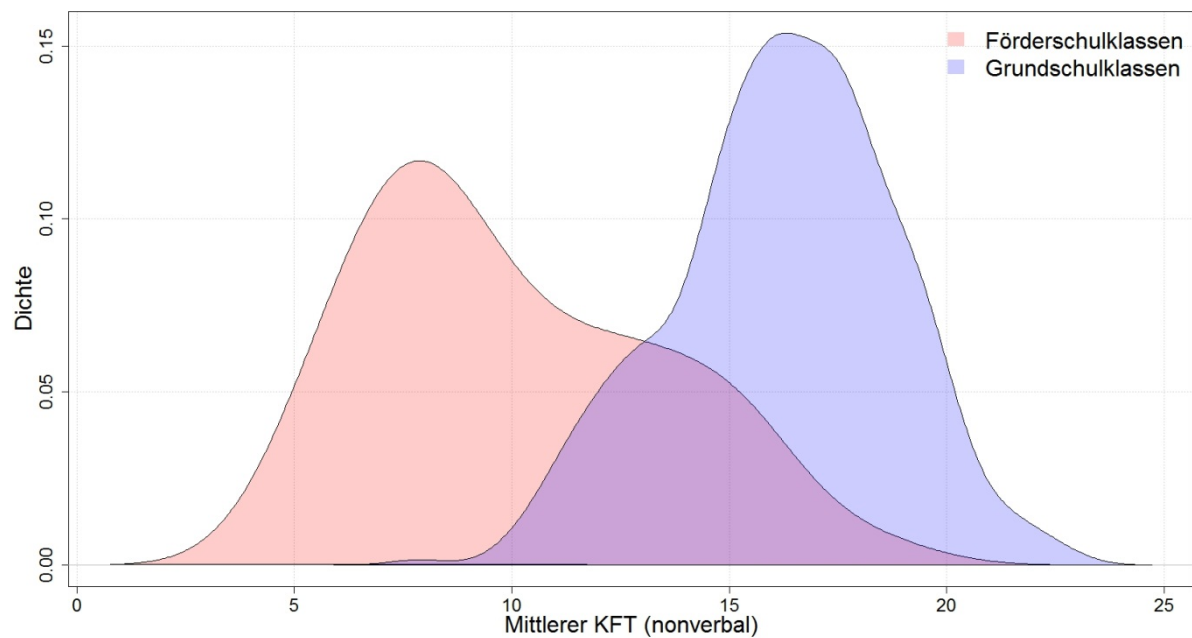
Ich danke Dr. Poldi Kuhl, Prof. Dr. Petra Stanat und Prof. Dr. Hans Anand Pant für die Betreuung dieser Arbeit. Vielen Dank, dass ihr mir mit Rat, Unterstützung und Ermutigung zur Seite standet und immer ein offenes Ohr für meine Fragen hattet! Liebe Poldi, liebe Petra, eure inhaltliche, organisatorische und motivationale Unterstützung war fantastisch und ist das Fundament dieser Arbeit. Prof. Dr. Manuel Völkle, Prof. Dr. Petra Stanat und Prof. Dr. Kai S. Cortina danke ich für ihre Bereitschaft zur Begutachtung der Dissertation. Mein Dank gilt darüber hinaus Prof. Dr. Matthias Ziegler und Dr. Martin Hecht für ihre Teilnahme an der Promotionskommission. Ich danke allen Koautorinnen und Koautoren der vier Einzelbeiträge dieser Dissertation für die angenehme und hilfreiche Zusammenarbeit: Dr. Nicole Haag, Dr. Malte Jansen, Dr. Anna Kroth, Dr. Poldi Kuhl, Prof. Dr. Hans Anand Pant, Dr. Camilla Rjosk, Prof. Dr. Petra Stanat und Dr. Sebastian Weirich. Sebastian Weirich und Nicole Haag danke ich für die tolle Einführung in Item Response Modelle mit dem R-Paket eatRep und ihre methodischen Ratschläge im Allgemeinen. Malte Jansen und Camilla Rjosk danke ich für ihre großartigen Nachhilfestunden in Mplus und den Diskussionen zu Mehrebenenmodellen. Mein Dank gilt auch der International Max Planck Research School on the Life Courses (LIFE), die mir die Möglichkeit gab, mich in einem interdisziplinären und internationalen Umfeld mit anderen Doktorandinnen und Doktoranden auszutauschen sowie von den interessanten Seminaren und Akademien zu profitieren. Außerdem danke ich meinen Kolleginnen und Kollegen am Forschungsdatenzentrum (FDZ) des IQB sowie des Projekts Forschungsdaten Bildung für die schöne Zusammenarbeit in den letzten Jahren. Ich danke dem FDZ am IQB auch für die Bereitstellung der Daten des IQB-Ländervergleichs 2011. Ich bedanke mich außerdem bei allen Kolleginnen und Kollegen des IQB – wegen euch freue ich mich (fast) jeden Morgen auf den Arbeitstag. Schließlich möchte ich mich bei meiner Familie (vielen Dank für das flinke und gründliche Lektorat, liebe Mama!) und meinen Freunden für die Unterstützung und Zerstreuung danken.

Anhang

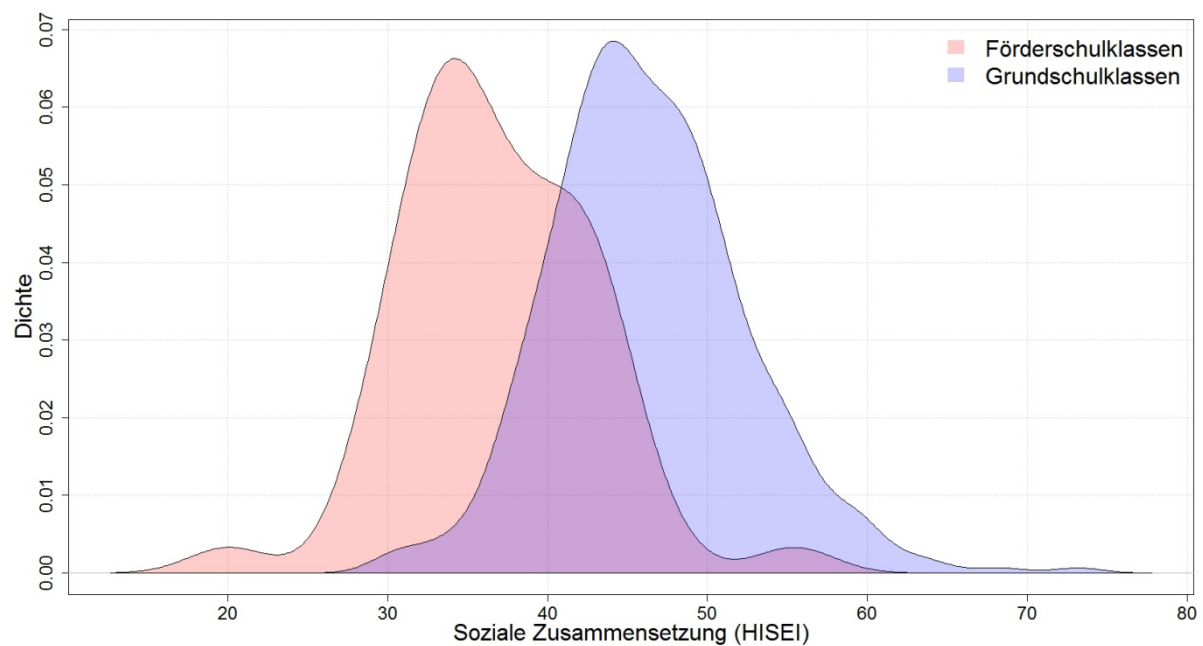
Inhaltsverzeichnis

- | | |
|----|--|
| A | Anhang zur Teilstudie 3: Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf |
| A1 | Anhang A1. Dichteverteilung der mittleren nonverbalen kognitiven Grundfähigkeiten (KFT) in Förder- und Grundschulklassen |
| A2 | Anhang A2. Dichteverteilung der sozialen Zusammensetzung (HISEI) in Förder- und Grundschulklassen. |
| B | Anhang zur Teilstudie 4: Educational placement and achievement motivation of students with special educational needs |
| B1 | Anhang B1. Testing for Measurement Invariance and Predicting Enjoyment of Learning With Mathematics Achievement |
| B2 | Anhang B2. Testing Measurement Invariance for Academic Self-Concept German |
| B3 | Anhang B3. Testing Measurement Invariance for Academic Self-Concept Mathematics |
| B4 | Anhang B4. Testing Measurement Invariance for Enjoyment of Learning |
| B5 | Anhang B5. Predicting SEN Students' Enjoyment of Learning With Mathematics Achievement |
| B6 | Anhang B6. Testing for Differential Effects of Type of Special Educational Needs on Students' Achievement Motivation |
| B7 | Anhang B7. Testing for Differential Effects of Type of SEN on Academic Self-Concept |
| B8 | Anhang B8. Testing for Differential Effects of Type of SEN on Enjoyment of Learning |

A. Anhang zur Teilstudie 3: Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf



Anhang A1. Dichteverteilung der mittleren nonverbalen kognitiven Grundfähigkeiten (KFT) in Förder- und Grundschulklassen.



Anhang A2. Dichteverteilung der sozialen Zusammensetzung (HISEI) in Förder- und Grundschulklassen.

B. Anhang zur Teilstudie 4: Educational placement and achievement motivation of students with special educational needs

Anhang B1. Testing for Measurement Invariance and Predicting Enjoyment of Learning With Mathematics Achievement

In supplementary analyses, we first checked the fit of the measurement model for each motivational outcome separately and second, we tested for measurement invariance between students with special educational needs (SEN) in special education schools and regular schools. Comparable measurement models and measurement invariance are necessary requirements for valid comparisons of the latent factor means across both student groups. First, for SEN students in each school setting (special education school vs. regular school), a two-level measurement model with equal factor loadings on the student level (L1) and on the classroom level (L2) was specified (see Model 0 in Anhang B2 - B4). Second, we ran different models constraining (1) all factor loadings to be equal for SEN students in special education schools and regular schools (metric invariance, see Model 1 in Anhang B2 - B4), (2) all intercepts to be equal across both groups (intercept-only invariance, see Model 2 in Anhang B2 - B4), and (3) allow one intercept to vary between groups while constraining the other intercepts and all factor loadings to be equal (partial invariance, see Model 3 in Anhang B2 - B4). The measurement models showed a reasonable fit for SEN students in special education schools and regular schools (see Model 0 in Anhang B2 - B4). Furthermore, we established partial invariance for each motivational outcome allowing valid inferences about differences in the latent means between groups. Anhang B5 shows that analyses with achievement in mathematics as predictor of students' enjoyment of learning yielded results similar to reading achievement as predictor of students' enjoyment of learning.

Anhang B2. Testing Measurement Invariance for Academic Self-Concept German

	χ^2 (df)	p	CFI	TLI	RMSEA	BIC	AIC	$\Delta\chi^2$ (df)	$\frac{p}{(\Delta\chi^2)}$
Model 0 ^a : Configural invariance (regular school)	0.90 (2)	.64	1.00	1.00	.00	4084.38	4028.54	-	-
Model 0 ^a : Configural invariance (special education school)	9.23 (2)	.01	.96	.89	.11	2573.14	2524.19	-	-
Model 1: Metric invariance	3.49 (3)	.32	1.00	1.00	.02	6610.31	6538.94	-	-
Model 2: Intercept-only invariance	26.36 (3)	.00	.95	.90	.14	6633.18	6561.81	22.87 (0)	.00
Model 3 ^b : Partial invariance	30.61 (8)	.00	.95	.96	.08	6603.63	6556.05	4.24 (5)	.52

^aModel 0 refers to a two-level measurement model with equal factor loadings on student level (L1) and classroom level (L2), ^bModel 3 allows differences in the intercept of item 1 between special education and regular schools and restricts residual variances to be equal across both groups

Anhang B3. Testing Measurement Invariance for Academic Self-Concept Mathematics

	χ^2 (df)	p	CFI	TLI	RMSEA	BIC	AIC	$\Delta\chi^2$ (df)	$\frac{p}{(\Delta\chi^2)}$
Model 0 ^a : Configural invariance (regular school)	2.24 (2)	.33	1.00	1.00	.02	4225.63	4169.75	-	-
Model 0 ^a : Configural invariance (special education school)	1.88 (2)	.39	1.00	1.00	.00	2497.57	2448.58	-	-
Model 1: Metric invariance	0.48 (3)	.92	1.00	1.01	.00	6686.29	6614.87	-	-
Model 2: Intercept-only invariance	19.75 (3)	.00	0.98	0.96	.11	6705.56	6634.14	19.27 (0)	.00
Model 3 ^b : Partial invariance	15.91 (5)	.01	0.99	0.98	.07	6688.20	6626.30	3.84 (2)	.15

^aModel 0 refers to a two-level measurement model with equal factor loadings on student level (L1) and classroom level (L2), ^bModel 3 allows differences in the intercept of item 1 between special education and regular schools

Anhang B4. Testing Measurement Invariance for Enjoyment of Learning

	χ^2 (df)	p	CFI	TLI	RMSEA	BIC	AIC	$\Delta\chi^2$ (df)	$\frac{p}{(\Delta\chi^2)}$
Model 0 ^a : Configural invariance (regular school)	3.81 (2)	.15	.99	.98	.04	4168.72	4112.91	-	-
Model 0 ^a : Configural invariance (special education school)	0.70 (2)	.71	1.00	1.01	.00	2526.17	2477.15	-	-
Model 1: Metric invariance	4.02 (3)	.26	1.00	1.00	.03	6661.89	6590.50	-	-
Model 2: Intercept-only invariance	19.43 (3)	.00	.97	.94	.11	6677.30	6605.91	15.41 (0)	.00
Model 3 ^b : Partial invariance	13.97 (5)	.02	.98	.98	.07	6658.32	6596.45	5.46 (2)	.07

^aModel 0 refers to a two-level measurement model with equal factor loadings on student level (L1) and classroom level (L2), ^bModel 3 allows differences in the intercept of item 3 between special education and regular schools

 Anhang B5. *Predicting SEN Students' Enjoyment of Learning With Mathematics Achievement*

	Enjoyment of learning									
	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)										
Mathematics achievement	0.09*	0.04	0.14**	0.05	0.13*	0.05			0.12**	0.05
Perceived social support							0.21***	0.03	0.21***	0.03
Gender (1 = female)	0.22**	0.07	0.22**	0.07	0.22**	0.07	0.14*	0.07	0.17*	0.07
Classroom level (L2)										
School type (1 = special education school)	0.30***	0.08			0.13	0.10	−0.24**	0.07	0.09	0.09
Class-average mathematics achievement			−0.19 ***	0.05	−0.13*	0.06			−0.16**	0.06
Class-average social support							0.06	0.04	0.08	0.04
Model fit										
Chi-Square test of model fit (df)	88.66*** (9)		82.45*** (9)		84.01*** (11)		33.83*** (11)		90.57*** (17)	
Bayesian Information Criteria (BIC)	9645.42		10582.22		9629.75		8992.37		11953.59	
Akaike Information Criteria (AIC)	9555.74		10482.24		9535.11		8901.55		11839.13	
Comparative Fit Index (CFI)	0.87		0.88		0.89		0.97		0.90	
Root Mean Square Error (RMSEA)	0.09		0.09		0.08		0.05		0.06	
Described Variance										
<i>R</i> ² on student level (L1)	4.30%		6.80%		6.30%		10.90%		14.10%	
<i>R</i> ² on classroom level (L2)	12.20%		21.10%		20.40%		12.70%		27.00%	

 SEN = special educational needs, Intraclass correlation (ICC) = .15., * $p < .05$. ** $p < .01$.

 *** $p < .001$

Anhang B6. Testing for Differential Effects of Type of Special Educational Needs on Students' Achievement Motivation

In supplementary analyses, we tested if the relation between educational placement and motivational outcomes differs between students with learning disabilities and students with speech or language impairment. The small amount of students with emotional disorders in special education schools prevented differential analyses of placement differences in academic self-concept and enjoyment of learning for this group. To test for differential effects of classroom composition on the relationship between type of SEN and academic self-concept and enjoyment of learning, respectively, random slope models were run in a first step. In the case of classrooms differing in the magnitude of the relation between type of SEN and both motivational outcomes (i.e., significant slope variation), we planned to apply cross-level-interaction to test whether those differences could be explained by classroom composition (Aguinis, Gottfredson, & Culpepper, 2013). After accounting for differences in individual achievement, no significant main effects of type of SEN were found (see Model 1 and Model 3 in Anhang B7 and Anhang B8). Random slope models were then estimated to test if the relation between type of SEN and motivational outcomes varied between classrooms. The slope variance was not statistically different from zero for both academic self-concept and enjoyment of learning (see Model 2 and Model 4 Anhang B7 and Anhang B8). Therefore, no additional cross-level interaction effects between class-average achievement and type of SEN were considered.

Anhang B7. *Testing for Differential Effects of Type of SEN on Academic Self-Concept*

	Academic self-concept German				Academic self-concept mathematics			
	Model 1		Model 2		Model 3		Model 4	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)								
Reading achievement	0.16***	0.04	0.16***	0.04				
Mathematics achievement					0.39***	0.05	0.39***	0.05
Type of SEN ¹ : Learning disabilities (1= yes)	−0.01	0.07	−0.01	0.07	−0.05	0.09	−0.06	0.09
Type of SEN ¹ : Speech or language impairment (1= yes)	−0.09	0.07	−0.09	0.07	0.11	0.08	0.12	0.09
Gender (1 = female)	0.12*	0.05	0.12*	0.05	−0.12	0.06	−0.12*	0.06
Classroom level (L2)								
School type (1 = special education school)	0.13	0.09	0.13	0.09	0.01	0.09	−0.02	0.09
Class-average reading achievement	−0.14*	0.06	−0.14*	0.07				
Class-average mathematics achievement					−0.35***	0.07	−0.36***	0.07
Variance Components								
Slope (L2) variance: Learning disabilities			0.04	0.27			0.12	0.22
Slope (L2) variance: Speech or language impairment			0.02	0.18			0.02	0.17
Intercept-slope (L2) covariance: Learning disabilities			−0.01	0.13			−0.001	0.12
Intercept-slope (L2) covariance: Speech or language impairment			−0.01	0.10			−0.02	0.10
Model fit								
Bayesian Information Criteria (BIC)	9677.06		9702.82		9582.05		9595.96	
Akaike Information Criteria (AIC)	9572.25		9578.04		9482.45		9476.44	

SEN = special educational needs, Intraclass correlation (ICC) = .09 for self-concept German and ICC = .14 for self-concept mathematics, * $p < .05$, ** $p < .01$, *** $p < .001$

Anhang B8. *Testing for Differential Effects of Type of SEN on Enjoyment of Learning*

	Enjoyment of learning (Reading achievement)				Enjoyment of learning (Mathematics achievement)			
	Model 1		Model 2		Model 3		Model 4	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Student level (L1)								
Reading achievement	0.14***	0.04	0.13**	0.04				
Mathematics achievement					0.15**	0.05	0.15**	0.05
Type of SEN ¹ : Learning disabilities (1= yes)	0.10	0.08	0.10	0.08	0.13	0.08	0.13	0.09
Type of SEN ¹ : Speech or language impairment (1= yes)	0.11	0.08	0.10	0.08	0.11	0.08	0.10	0.09
Gender (1 = female)	0.18**	0.06	0.18**	0.07	0.21**	0.07	0.21**	0.07
Classroom level (L2)								
School type (1 = special education school)	0.06	0.12	0.07	0.11	0.10	0.11	0.12	0.12
Class-average reading achievement	−0.15*	0.07	−0.16*	0.07				
Class-average mathematics achievement					−0.14*	0.07	−0.13*	0.07
Variance Components								
Slope (L2) variance: Learning disabilities			0.05	0.19			0.04	0.22
Slope (L2) variance: Speech or language impairment			0.10	0.21			0.11	0.21
Intercept-slope (L2) covariance: Learning disabilities			−0.002	0.10			−0.003	0.12
Intercept-slope (L2) covariance: Speech or language impairment			−0.08	0.12			−0.09	0.12
Model fit								
Bayesian Information Criteria (BIC)	9708.44		9728.01		9641.11		9661.29	
Akaike Information Criteria (AIC)	9603.64		9603.25		9536.51		9536.76	

SEN = special educational needs, Intraclass correlation (ICC) = .15, * $p < .05$, ** $p < .01$,*** $p < .001$